


Topology aggregation using parameter obtained by internodal negotiation

Patent Number: ☐ EP0876076, A3
Publication date: 1998-11-04
Inventor(s): IWATA ATSUSHI (JP)
Applicant(s):: NIPPON ELECTRIC CO (JP)
Requested Patent: ☒ JP11015804
Application Number: EP19980107852 19980429
Priority Number(s): JP19970126532 19970430; JP19970365607 19971222
IPC Classification: H04Q11/04 ; H04L12/56
EC Classification: H04L12/56E, H04Q3/00D4, H04Q11/04S2
Equivalents: CA2236384, JP3063721B2

Abstract

In a hierarchical network, an aggregation parameter is exchanged between a node and its neighbor to agree on a negotiated parameter. A set of physical links between them are aggregated into a logical link according to the negotiated parameter and resource data of the logical link is stored in a database of each node. The resource data is then exchanged between them and the contents of the database of each node are updated with resource data of its neighbor. To optimize the link aggregation, an operating performance of the logical link is evaluated by the node using traffic data from the network, and the negotiated parameter is updated according to the detected operating performance. The updated aggregation parameter of the node is exchanged with its neighbor to agree on a negotiated updated parameter which is used to aggregate the physical links into the logical link. 

Data supplied from the esp@cenet database - I2

THIS PAGE BLANK (USPTO)

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-15804

(43) 公開日 平成11年(1999) 1月22日

(51) Int.Cl.⁴

G 0 6 F 15/177

識別記号

F I

G 0 6 F 15/16

4 0 0 D

審査請求 有 請求項の数18 F D (全 32 頁)

(21) 出願番号 特願平9-365607

(22) 出願日 平成9年(1997)12月22日

(31) 優先権主張番号 特願平9-126532

(32) 優先日 平9(1997)4月30日

(33) 優先権主張国 日本 (J P)

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 岩田 淳

東京都港区芝五丁目7番1号 日本電気株式会社内

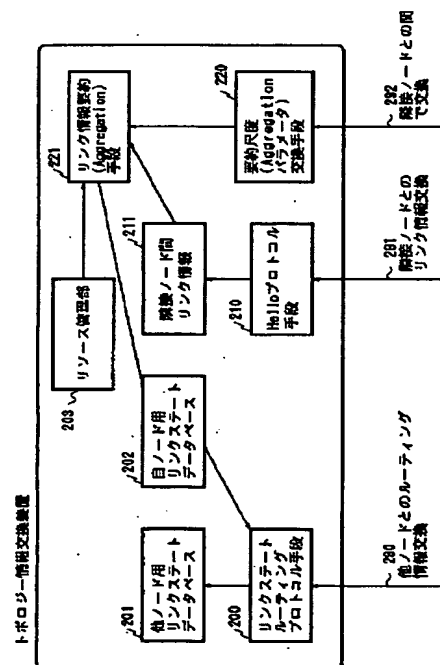
(74) 代理人 弁理士 境 廣巳

(54) 【発明の名称】 トポロジー情報交換装置及びプログラムを記録した機械読み取り可能な記録媒体

(57) 【要約】

【課題】 大規模な階層的なネットワークにおいて、各階層内のドメイン内部の、リンク、或いはノードのトポロジー情報を効果的に要約して、他のドメインへ提供する方式を実現する。

【解決手段】 ハロープロトコル手段210は隣接ノード間に存在する物理リンクを認識し、その情報を隣接ノード間リンク情報211に格納する。要約尺度交換手段220は、複数の物理リンクの異なるQOSを1本或いは少数の論理リンクに要約する場合に、QOSの値を過大評価と過少評価との間でどの程度の割合で要約するかのパラメータをお互いに交換する。リンク情報要約手段221は、上記のパラメータと隣接ノード間リンク情報211とから複数の物理リンクを要約し、そのリンク属性として新規のQOS値を計算する。



【特許請求の範囲】

【請求項1】 ノード間でトポロジー情報を交換するトポロジー情報交換装置において、隣接ノード間の複数リンク情報の要約尺度を隣接ノード間で交換する要約尺度交換手段と、当該要約尺度に基づきリンク情報を要約して計算するリンク情報要約手段と、ノード間で当該要約リンク情報をお互いに交換するトポロジー交換手段とから構成されることを特徴とするトポロジー情報交換装置。

【請求項2】 請求項1記載のトポロジー情報交換装置において、ネットワークの性能評価手段により得られた性能評価結果に基づき新たな要約尺度を求め直す要約尺度計算手段を備え、当該新規要約尺度を隣接ノード間で交換することにより、あらたにリンク情報を要約し直し、ノード間で新規要約リンク情報をお互いに交換する構成を有することを特徴とするトポロジー情報交換装置。

【請求項3】 ノード間でトポロジー情報を交換するトポロジー情報交換装置において、隣接ノード間の複数リンクを類似した品質ごとにグループ化するリンクグループ作成手段と、隣接ノード間の複数リンク情報の要約尺度をノード間で交換する要約尺度交換手段と、当該要約尺度に基づき各リンクグループ毎に要約リンク情報を計算するリンク情報要約手段と、ノード間で各リンクグループ毎の要約リンク情報をお互いに交換するトポロジー交換手段とから構成されることを特徴とするトポロジー情報交換装置。

【請求項4】 ノード間でトポロジー情報を交換するトポロジー情報交換装置において、ネットワークの性能評価手段と、該ネットワークの性能評価手段で性能評価結果が求められる毎にその性能評価結果に基づき隣接ノード間の複数リンクを再グループ化するリンクグループ作成手段と、隣接ノード間の複数リンク情報の要約尺度をノード間で交換する要約尺度交換手段と、前記再グループ化されたリンクグループ毎に要約リンク情報を計算し直すリンク情報要約手段と、ノード間で各リンクグループ毎の要約リンク情報をお互いに交換するトポロジー交換手段とから構成されることを特徴とするトポロジー情報交換装置。

【請求項5】 ノード間でトポロジー情報を交換するトポロジー情報交換装置において、ネットワークの性能評価手段と、該ネットワークの性能評価手段で性能評価結果が求められる毎にその性能評価結果に基づき隣接ノード間の複数リンク情報の要約尺度を求め直す要約尺度計算手段と、前記ネットワーク性能評価手段で性能評価結果が求められる毎にその性能評価結果に基づき隣接ノード間の複数リンクを再グループ化するリンクグループ作成手段と、前記要約尺度計算手段で求め直された要約尺度をノード間で交換する要約尺度交換手段と、当該要約尺度を用いて前記再グループ化されたリンクグループ毎

に要約リンク情報を計算し直すリンク情報要約手段と、ノード間で各リンクグループ毎の要約リンク情報をお互いに交換するトポロジー交換手段とから構成されることを特徴とするトポロジー情報交換装置。

【請求項6】 階層ネットワークにおいて上位ノードが自ノード配下の下位ドメインのトポロジー情報を代表的に交換するトポロジー情報交換装置において、当該下位ドメイン内のノードのうち他のドメインと接続している境界ノードを検出する境界ノード検出手段と、当該検出境界ノード間のトポロジーを所定のトポロジーテンプレートで表現し、実際のトポロジーとネットワークの品質とを前記トポロジーテンプレートにより近似表現するトポロジー要約手段と、ノード間で新規要約トポロジー情報をお互いに交換するトポロジー交換手段とから構成されることを特徴とするトポロジー情報交換装置。

【請求項7】 階層ネットワークにおいて上位ノードが自ノード配下の下位ドメインのトポロジー情報を代表的に交換するトポロジー情報交換装置において、当該下位ドメイン内のノードのうち他のドメインと接続している境界ノードを検出する境界ノード検出手段と、下位ドメイン間の複数リンクを一つあるいは少数の上位ノード間のリンクの要約に伴い、複数の前記境界ノードを要約する境界ノード要約手段と、当該要約境界ノード間のトポロジーを所定のトポロジーテンプレートで表現し、実際のトポロジーとネットワークの品質とを前記トポロジーテンプレートにより近似表現するトポロジー要約手段と、ノード間で新規要約トポロジー情報をお互いに交換するトポロジー交換手段とから構成されることを特徴とするトポロジー情報交換装置。

【請求項8】 階層ネットワークにおいて上位ノードが自ノード配下の下位ドメインのトポロジー情報を代表的に交換するトポロジー情報交換装置において、上位ノード間でトポロジー情報の要約尺度を交換する要約尺度交換手段と、下位ドメイン内のノードのうち他のドメインと接続している境界ノードを検出する境界ノード検出手段と、下位ドメイン間の複数リンクを一つあるいは少数の上位ノード間のリンクの要約に伴い、複数の前記境界ノードを要約する境界ノード要約手段と、前記要約尺度に基づき前記要約境界ノード間のトポロジーを所定のトポロジーテンプレートで表現し、実際のトポロジーとネットワークの品質とを前記トポロジーテンプレートにより近似表現するトポロジー要約手段と、ノード間で新規要約トポロジー情報をお互いに交換するトポロジー交換手段とから構成されることを特徴とするトポロジー情報交換装置。

【請求項9】 階層ネットワークにおいて上位ノードが自ノード配下の下位ドメインのトポロジー情報を代表的に交換するトポロジー情報交換装置において、ネットワークの性能評価手段と、該ネットワークの性能評価手段で性能評価結果が求められる毎にその性能評価結果に基

づきトポロジー情報の要約尺度を求め直す要約尺度計算手段と、該求められた要約尺度を上位ノード間で交換する要約尺度交換手段と、下位ドメイン内のノードのうち他のドメインと接続している境界ノードを検出する境界ノード検出手段と、下位ドメイン間の複数リンクを一つあるいは少数の上位ノード間のリンクの要約に伴い、複数の前記境界ノードを要約する境界ノード要約手段と、前記要約尺度に基づき前記要約境界ノード間のトポロジーを所定のトポロジーテンプレートで表現し、実際のトポロジーとネットワークの品質とを前記トポロジーテンプレートにより近似表現するトポロジー要約手段と、ノード間で新規要約トポロジー情報をお互いに交換するトポロジー交換手段とから構成されることを特徴とするトポロジー情報交換装置。

【請求項10】 ノード間でトポロジー情報を交換するトポロジー情報交換装置を構成するコンピュータを、隣接ノード間の複数リンク情報の要約尺度を隣接ノード間で交換する要約尺度交換手段、当該要約尺度に基づきリンク情報を要約して計算するリンク情報要約手段、ノード間で当該要約リンク情報をお互いに交換するトポロジー交換手段、として機能させるプログラムを記録した機械読み取り可能な記録媒体。

【請求項11】 ノード間でトポロジー情報を交換するトポロジー情報交換装置を構成するコンピュータを、ネットワークの性能評価手段、該ネットワークの性能評価手段で性能評価結果が求められる毎にその性能評価結果に基づき隣接ノード間の複数リンク情報の要約尺度を求め直す要約尺度計算手段、新たに求められた要約尺度を隣接ノード間で交換する要約尺度交換手段、当該要約尺度に基づきリンク情報を要約し直すリンク情報要約手段、ノード間で当該要約リンク情報をお互いに交換するトポロジー交換手段、として機能させるプログラムを記録した機械読み取り可能な記録媒体。

【請求項12】 ノード間でトポロジー情報を交換するトポロジー情報交換装置を構成するコンピュータを、隣接ノード間の複数リンクを類似した品質ごとにグループ化するリンクグループ作成手段、隣接ノード間の複数リンク情報の要約尺度をノード間で交換する要約尺度交換手段、各リンクグループ毎に要約リンク情報を計算するリンク情報要約手段、ノード間で各リンクグループ毎の要約リンク情報をお互いに交換するトポロジー交換手段、として機能させるプログラムを記録した機械読み取り可能な記録媒体。

【請求項13】 ノード間でトポロジー情報を交換するトポロジー情報交換装置を構成するコンピュータを、ネットワークの性能評価手段、該ネットワークの性能評価手段で性能評価結果が求められる毎にその性能評価結果に基づき隣接ノード間の複数リンクを再グループ化するリンクグループ作成手段、隣接ノード間の複数リンク情報の要約尺度をノード間で交換する要約尺度交換手段、

前記再グループ化されたリンクグループ毎に要約リンク情報を計算し直すリンク情報要約手段、ノード間で各リンクグループ毎の要約リンク情報をお互いに交換するトポロジー交換手段、として機能させるプログラムを記録した機械読み取り可能な記録媒体。

【請求項14】 ノード間でトポロジー情報を交換するトポロジー情報交換装置を構成するコンピュータを、ネットワークの性能評価手段、該ネットワークの性能評価手段で性能評価結果が求められる毎にその性能評価結果に基づき隣接ノード間の複数リンク情報の要約尺度を求め直す要約尺度計算手段、前記ネットワーク性能評価手段で性能評価結果が求められる毎にその性能評価結果に基づき隣接ノード間の複数リンクを再グループ化するリンクグループ作成手段、前記要約尺度計算手段で求め直された要約尺度をノード間で交換する要約尺度交換手段、当該要約尺度を用いて前記再グループ化されたリンクグループ毎に要約リンク情報を計算し直すリンク情報要約手段、ノード間で各リンクグループ毎の要約リンク情報をお互いに交換するトポロジー交換手段、として機能させるプログラムを記録した機械読み取り可能な記録媒体。

【請求項15】 階層ネットワークにおいて上位ノードが自ノード配下の下位ドメインのトポロジー情報を代表的に交換するトポロジー情報交換装置を構成するコンピュータを、当該下位ドメイン内のノードのうち他のドメインと接続している境界ノードを検出する境界ノード検出手段、当該検出境界ノード間のトポロジーを所定のトポロジーテンプレートで表現し、実際のトポロジーとネットワークの品質とを前記トポロジーテンプレートにより近似表現するトポロジー要約手段、ノード間で新規要約トポロジー情報をお互いに交換するトポロジー交換手段、として機能させるプログラムを記録した機械読み取り可能な記録媒体。

【請求項16】 階層ネットワークにおいて上位ノードが自ノード配下の下位ドメインのトポロジー情報を代表的に交換するトポロジー情報交換装置を構成するコンピュータを、当該下位ドメイン内のノードのうち他のドメインと接続している境界ノードを検出する境界ノード検出手段、下位ドメイン間の複数リンクを一つあるいは少数の上位ノード間のリンクの要約に伴い、複数の前記境界ノードを要約する境界ノード要約手段、当該要約境界ノード間のトポロジーを所定のトポロジーテンプレートで表現し、実際のトポロジーとネットワークの品質とを前記トポロジーテンプレートにより近似表現するトポロジー要約手段、ノード間で新規要約トポロジー情報をお互いに交換するトポロジー交換手段、として機能させるプログラムを記録した機械読み取り可能な記録媒体。

【請求項17】 階層ネットワークにおいて上位ノードが自ノード配下の下位ドメインのトポロジー情報を代表的に交換するトポロジー情報交換装置を構成するコンピ

10

20

30

40

50

ュータを、上位ノード間でトポロジー情報の要約尺度を交換する要約尺度交換手段、下位ドメイン内のノードのうち他のドメインと接続している境界ノードを検出する境界ノード検出手段、下位ドメイン間の複数リンクを一つあるいは少数の上位ノード間のリンクの要約に伴い、複数の前記境界ノードを要約する境界ノード要約手段、前記要約尺度に基づき前記要約境界ノード間のトポロジーを所定のトポロジーテンプレートで表現し、実際のトポロジーとネットワークの品質とを前記トポロジーテンプレートにより近似表現するトポロジー要約手段、ノード間で新規要約トポロジー情報をお互いに交換するトポロジー交換手段、として機能させるプログラムを記録した機械読み取り可能な記録媒体。

【請求項18】 階層ネットワークにおいて上位ノードが自ノード配下の下位ドメインのトポロジー情報を代表的に交換するトポロジー情報交換装置を構成するコンピュータを、ネットワークの性能評価手段、該ネットワークの性能評価手段で性能評価結果が求められる毎にその性能評価結果に基づきトポロジー情報の要約尺度を求め直す要約尺度計算手段、該求められた要約尺度を上位ノード間で交換する要約尺度交換手段、下位ドメイン内のノードのうち他のドメインと接続している境界ノードを検出する境界ノード検出手段、下位ドメイン間の複数リンクを一つあるいは少数の上位ノード間のリンクの要約に伴い、複数の前記境界ノードを要約する境界ノード要約手段、前記要約尺度に基づき前記要約境界ノード間のトポロジーを所定のトポロジーテンプレートで表現し、実際のトポロジーとネットワークの品質とを前記トポロジーテンプレートにより近似表現するトポロジー要約手段、ノード間で新規要約トポロジー情報をお互いに交換するトポロジー交換手段、として機能させるプログラムを記録した機械読み取り可能な記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、ネットワークにおけるノード間でリンク情報を経路情報としてお互いに交換するトポロジー情報交換装置、および、大規模な階層的なネットワークにおいて、上位階層のノード間で自ノード配下の下位ドメインのトポロジー情報を経路情報として交換するトポロジー情報交換装置に関し、特にトポロジー情報を効果的に要約(Aggregation)することにより、ネットワーク全体としての経路情報交換量の削減、コネクション設定失敗(ブロック)率の低減を実現するトポロジー情報交換装置に関する。

【0002】

【従来の技術】ネットワークにおけるノード間で、各ノードが自分に接続されているリンクの情報(リンクのトポロジー情報)を他の全てのノードへ通知し、各ノードがその情報を全て受信することにより、全てのノードにおいて、ノードとリンクの接続関係が全て分かり、ネッ

トワーク全体のトポロジーを認識することができる。また、大規模な階層的なネットワークにおいては、上位階層のノード間で自ノード配下の下位ドメインのトポロジー情報、つまりそのドメイン内に存在するリンク、ノードのトポロジー情報を交換することにより、上位階層において下位階層のトポロジーを認識することができる。

【0003】トポロジー情報の要約は、経路情報交換量の削減、コネクション設定失敗(ブロック)率の低減を実現することを目的として行われる。トポロジー情報の要約に関して記載した文献として、1996年3月発行の「ザ エイティエム フォーラム プライベート ネットワーク ネットワーク インタフェース スペシフィケーション バージョン 1.0 (The ATM Forum Private Network-Network Interface Specification Version 1.0)がある。この文献には、トポロジーの要約の表現方法(テンプレート)に関する規定がされているが、どのような形でトポロジー情報を要約するかについての規定はない。このため、具体的な要約方式についてはベンダーが自由に選択することが可能であるが、具体的な要約方式を記載した文献は現在のところ見当たらない。

【0004】

【発明が解決しようとする課題】上述したように従来においては、要約のための表現用のテンプレートは規定されていたが、具体的な要約方式については規定されておらず、ベンダーの自由な選択に任されている。このため、トポロジー情報を効果的に要約し得る技術の開発が望まれている。

【0005】そこで本発明の目的は、ネットワークにおけるノード間でリンク情報を効果的に要約することにより、ネットワーク全体としての経路情報交換量の削減、ならびにコネクション設定失敗(ブロック)率の低減を実現することにある。

【0006】また本発明の別の目的は、大規模な階層的なネットワークにおいて、上位階層のノード間で自ノード配下の下位ドメインのトポロジー情報を効果的に要約することにより、ネットワーク全体としての経路情報交換量の削減、ならびにコネクション設定失敗(ブロック)率の低減を実現することにある。

【0007】

【課題を解決するための手段】第1の発明は、ノード間でトポロジー情報を交換するトポロジー情報交換装置(図1の110、111、112)において、隣接ノード間の複数リンク情報の要約尺度を隣接ノード間で交換する要約尺度交換手段(図2の220)と、当該要約尺度に基づきリンク情報を要約して計算するリンク情報要約手段(図2の221)と、ノード間で当該要約リンク情報をお互いに交換するトポロジー交換手段(図2の200)とから構成される。ここで、要約尺度とは、複数

のリンクの異なるQOSを1本のリンク或いは少数のリンクに要約する場合に、QOSの値をAggressiveなAggregation（過大評価）とConservativeなAggregation（過小評価）との間でどの程度の割合で、要約を行うかのを示す尺度である。

【0008】従来はどのような形でトポロジー情報を要約するかについての規定はなく、ましてや要約尺度というものはノード間にはない。要約尺度を隣接ノード間で交換せず、各ノードが勝手に要約すると、隣接ノードA-B間の或るリンクに対して、同じリンク情報をA側は過大に評価し、B側は過小に評価することが起こり、同じリンクでありながら、表現精度がバラバラになってしまう。その結果、ノードA→Bへの通信の場合はリソースが多いという情報がきていることから比較的その経路を選択し易いのに、逆にノードB→Aの方はリソースがほとんどないという通知からその経路は選択されにくいという結果になる。通常、通信は双方向通信が一般的なので、A→BはOKなのに、B→Aの経路はだめということからその経路を使えなくなる可能性が高くなる。第1の発明では、隣接ノード間の複数リンク情報の要約尺度を隣接ノード間で交換し、その要約尺度に基づきリンク情報を要約して交換するため、A側もB側も同じ尺度で要約でき、上記のような問題が極力防止される。

【0009】第2の発明は、第1の発明のトポロジー情報交換装置において、ネットワークの性能評価手段（図11の700）により得られた性能評価結果に基づき、要約尺度計算手段（図11の701）が新たな要約尺度を求め直し、当該新規要約尺度を隣接ノード間で交換する（図11の292）ことにより、あらたにリンク情報を要約し直し、ノード間で新規要約リンク情報をお互いに交換する（図11の290）。ここで、ネットワークの性能評価項目としては、呼損率等を用いることができる。

【0010】要約尺度が適切でないと、隣接ノード間で同じ尺度で要約しても経路が選択されにくくなるような無駄な場合が起こり得る。例えば過小評価し過ぎると、ネットワークの帯域が無駄に余ってしまうといったネットワーク性能の低下が生じる。第2の発明では、そのような場合にネットワーク性能評価結果に基づいて要約尺度が動的に変更される。これにより、要約尺度の最適化が図られる。

【0011】第3の発明は、ノード間でトポロジー情報を交換するトポロジー情報交換装置において、隣接ノード間の複数リンクを類似した品質ごとにグループ化するリンクグループ作成手段（図14の901）と、隣接ノード間の複数リンク情報の要約尺度をノード間で交換する要約尺度交換手段（図14の220）と、当該要約尺度に基づき各リンクグループ毎に要約リンク情報を計算するリンク情報要約手段（図14の902）と、ノード

間で各リンクグループ毎の要約リンク情報をお互いに交換するトポロジー交換手段（図14の200）とから構成される。

【0012】例えば極端な例として、帯域が0で遅延が1msecのリンクと、帯域が100Mbpsで遅延が1secのリンクとの2本のリンクがあると、これを1本のリンクに要約すると、過少評価による要約では、要約リンクの帯域は0、遅延は1secとなり、過大評価による要約では、要約リンクの帯域は100Mbps、遅延は1msecとなる。過少評価による要約ではリンクとして全く意味のない品質が悪すぎるリンクとなり、過大評価による要約では非常に品質は良いが現実にはそんなリンクは存在しないという状況になる。また、過大評価では、無駄なコネクションがこのようなリンクを通過する可能性が増え、呼損率が増えてしまう。つまり、いずれも実際のリンク特性からはほど遠いという状況であり、このように特性の異なり過ぎるリンクを無理に1つのリンクに要約するのは却って良くない状況を引き起こす。したがって、リンク特性があまり類似していない場合には、要約せずにそれぞれ別々に取り扱う方が、より正確にリンクを記述することができる。つまり、要約することによって却って悪い状況を引き起こす場合は要約しない方が良いという観点から、第3の発明ではリンクグループという概念を導入している。

【0013】第4の発明は、第3の発明のトポロジー情報交換装置において、ネットワークの性能評価手段（図18の700）により得られた性能評価結果に基づき、リンクグループの再構成を行い（図18の901）、再構成したリンクグループ毎にあらたに要約リンク情報を計算し直し（図18の902）、ノード間で新規要約リンク情報をお互いに交換する（図18の290）。

【0014】リンクグループが適切でないと、リンクグループ毎の要約リンク情報も適切でなくなり、その結果として呼損率が増えたり、ネットワーク帯域が無駄に余ってしまう等のネットワーク性能の低下が生じる。第4の発明では、そのような場合にネットワーク性能評価結果に基づきリンクグループの再構成を行う。これにより、リンクグループの最適化が図られる。

【0015】第5の発明は第3の発明のトポロジー情報交換装置において、ネットワークの性能評価手段（図18の700）により得られた性能評価結果に基づき、リンクグループの再構成を行うと共に（図18の901）、要約尺度を求め直して（図18の701）ノード間で交換する（図18の292）ことにより、あらたに要約リンク情報を計算し直し、ノード間で新規要約リンク情報をお互いに交換する（図18の290）。

【0016】第5の発明によれば、第4の発明による効果に加えて、第2の発明と同様に要約尺度の最適化が図れる。

10

20

30

40

50

【0017】第6の発明は、階層ネットワークにおいて上位ノードが自ノード配下の下位ドメインのトポロジー情報を代表的に交換するトポロジー情報交換装置において、前記下位ドメイン内のノードのうち他のドメインと接続している境界ノードを検出する境界ノード検出手段（図23の1520）と、当該検出境界ノード間のトポロジーを所定のトポロジーテンプレートで表現し、実際のトポロジーとネットワークの品質とを前記トポロジーテンプレートにより近似表現するトポロジー要約手段（図23の1522）とから構成され、ノード間で新規

要約トポロジー情報をお互いに交換する（図23の1590）。ここで、所定のトポロジーテンプレートとしては、Symmetric Star、Asymmetric Star、Exception等を用いることができる。

【0018】大規模な階層的なネットワークにおいて、ネットワーク全体のトポロジー情報を認識し、経路選択を行う場合、特にあるドメイン（Peer Group）のリンク及びノードのトポロジー情報を全てのドメインに提供すると、経路情報が過多となり、大規模網に提供できない。第6の発明によれば、効果的に内部のトポロジー情報を要約することができ、ネットワーク全体としての経路情報交換量の削減が図れる。また、経路情報を削減することだけが目的であれば、あるドメインの内部の経路情報を全く出さない選択ができる。しかしながら、その場合、そのドメインを通過するトラヒックにとって、そのドメインを通過すべきか、他のドメインを通過した方が最適かを計算することができない。第6の発明では、効果的に内部のトポロジー情報を要約して交換することにより、ドメイン内部のある程度の荒いトポロジー情報、あるいはQOS情報を知ることができるため、ドメイン間の経路選択において、どのドメインを経由すべきという最適経路選択ができるかぎり可能となり、コネクション設定失敗（ブロック）率等を低減することができる。

【0019】第7の発明は、階層ネットワークにおいて上位ノードが自ノード配下の下位ドメインのトポロジー情報を代表的に交換するトポロジー情報交換装置において、当該下位ドメイン内のノードのうち他のドメインと接続している境界ノードを検出する境界ノード検出手段（図23の1520）と、下位ドメイン間の複数リンクの一つあるいは少数の上位ノード間のリンクの要約に伴い、複数の前記境界ノードを要約する境界ノード要約手段（図23の1523）と、当該要約境界ノード間のトポロジーを所定のトポロジーテンプレートで表現し、実際のトポロジーとネットワークの品質とを当該トポロジーテンプレートにより近似表現するトポロジー要約手段（図23の1522）とから構成され、ノード間で新規要約トポロジー情報をお互いに交換する（図23の1590）。

【0020】階層的な経路情報はすべてのノードが同じものを持つように制御が加えられるため、大規模ネットワークをサポートするためには、なるべく情報量を削減する必要がある。第7の発明によれば、上位レイヤの論理ノード間の論理リンクの要約に伴って境界ノードも要約することにより、経路情報量のより一層の削減が図れる。

【0021】第8の発明は、第6の発明のトポロジー情報交換装置において、上位ノード間でトポロジー情報の要約尺度を交換する要約尺度交換手段（図27の1700）によって得られた要約尺度に基づき、新規に当該下位ドメインのトポロジー情報を計算し直し（図27の1522）、上位ノード間で当該新規要約トポロジー情報をお互いに交換する（図27の1590）。ここで、トポロジー情報の要約尺度とは、検出境界ノード間のトポロジーを所定のトポロジーテンプレートで表現し、実際のトポロジーとネットワークの品質とを前記トポロジーテンプレートにより近似表現する場合に、Aggressiveに行うか、Conservativeに行うかを示すパラメータ、近似計算式における収束条件を示すパラメータのことである。

【0022】要約尺度を上位ノード間で交換せず、各上位ノードが勝手に要約すると、各ドメインの表現精度があいまいになってしまう。従って、或る下位ドメインAは過大評価され過ぎており、別の下位ドメインBは過少評価され過ぎて起ることが起こり、ドメインAにトラフィックが集中する結果、コネクション設定失敗率等が高まる弊害がある。第8の発明では、上位論理ノード間で要約尺度を交換し、その要約尺度に基づき各ドメインのトポロジー情報を要約して交換するため、そのような問題が極力防止される。

【0023】第9の発明は、第8の発明のトポロジー情報交換装置において、ネットワークの性能評価手段（図29の1800）により得られた性能評価結果に基づき、要約尺度計算手段（図29の1801）が要約尺度を求め直し、当該新規要約尺度を上位ノード間で交換することにより、あらたに下位ドメインのトポロジー情報を計算し直し（図29の1522）、上位ノード間で当該新規要約トポロジー情報をお互いに交換する（図29の1590）。

【0024】この第9の発明によれば、性能評価結果に基づき要約尺度を求め直すことで、要約尺度の最適化が図られる。

【0025】

【発明の実施の形態】次に本発明の実施の形態の例について図面を参照して詳細に説明する。

【0026】図1は本発明のトポロジー情報交換装置を適用したネットワークの簡単な例を示す。図1のネットワークにおいては、ノードとしてATMノード100、101、102が、物理リンク130、131、132

によって相互に接続されている。ここで各ATMノード100, 101, 102は、ATMスイッチのハードウェア120, 121, 122と、データの経路制御を行うためのトポロジー情報を交換するための本発明にかかるトポロジー情報交換装置110, 111, 112とから構成される。トポロジー情報交換装置110, 111, 112は、お互いに制御情報をやりとりすることにより、例えばATMノード100, 101間のリンク130の3本の物理リンクを認識し、他のATMノード102に対してトポロジー情報を配布する時に、1本の論理リンクに要約して配布する制御を行う。

【0027】図2は第1の発明にかかるトポロジー情報交換装置の実施の形態の機能構成を示し、図1内におけるトポロジー情報交換装置110, 111, 112内の機能構成を示している。図2に示したトポロジー情報交換装置は、リンクステートルーティングプロトコル手段200、他ノード用リンクステートデータベース201、自ノード用リンクステートデータベース202、リソース管理部203、Helloプロトコル手段210、隣接ノード間リンク情報211、要約尺度(Aggregationパラメータ)交換手段220およびリンク情報要約(Aggregation)手段221の8つの機能手段から構成されている。なお、図2の装置は、例えばCPU、メモリ等を有するコンピュータと制御プログラムを記録した記録媒体(磁気ディスクや半導体メモリ等)とで実現され、記録媒体中の制御プログラムがコンピュータに読み込まれ、コンピュータの動作を制御し、コンピュータを上記8つの機能手段として機能させる。

【0028】ハロー(Hello)プロトコル手段210は、隣接ノード間の複数の全ての物理リンクにおいて制御メッセージのやりとりを行い、隣接ノード間に存在する物理リンクを認識し、隣接ノード間リンク情報211にその情報を格納する。

【0029】図3にHello Packetでやりとりされる情報の例を示す。必須の情報は、自ノードID、自ポートID、相手ノードID、相手ポートID、自ドメインID(Peer Group ID)の5つである(これ以外の情報も標準のATM Forum Specificationでは存在するが、ここでは省略している)。Hello Packetによる制御メッセージのやりとりは、以下のように行われる。

【0030】(A) 自Peer Group IDを交換することで、隣接ノードが同じPeer Groupに属するかどうかを判断する。同じPeer Groupでなければより上位階層で接続されることになる。

(B) 同じPeer Groupと認識すれば、それぞれが自ノードID、自ポートIDのみの値を指定して相手に通知することにより、両者は相手側のノードID、ポートIDを認識できる。それにより、自分のノードの

どのポートが相手のノードのどのポートに接続されているかという情報を取得できる。

(C) 以後、自分の認識している自ノードID、自ポートID、相手ノードID、相手ポートID、自ドメインID(Peer Group ID)情報をお互いに定期的に交換することで、状態確認を絶えず行う。

【0031】隣接ノード間リンク情報211には、Helloプロトコル手段210によって交換された情報すべてが格納される。例えば隣接ノード間に複数の物理リンクがある場合の例を挙げると、図4のようになる。これには物理リンクのリソース情報(帯域、遅延など)は含まれない。各物理リンクの帯域、遅延などのリソース使用量、空き容量は、リソース管理部203で管理されている。

【0032】一方、要約尺度(アグリゲーションパラメータ: Aggregation パラメータ)交換手段220は、複数の物理リンクの異なるQOS(Quality of Service)を1本の論理リンク或いは少数の論理リンクに要約する場合に、QOSの値をAggressiveなAggregation(過大評価)とConservativeなAggregation(過小評価)との間でどの程度の割合で、Aggregationを行うかのパラメータ r ($0 \leq r \leq 1$, Aggressive=1, Conservative=0)をお互いに交換する。

【0033】図5に、Aggregationパラメータ手段220が隣接ノード間で行うパラメータのNegotiationの方法の例について示す。この例では、ノードAに対してParameter=1のデフォルト値が設定され、ノードBに対してParameter=0.5のデフォルト値が設定されている状況を想定しており、ノードAからParameter=1で申告したのに対して(フロー1200)、ノードBからParameter=0.5での申告があり(フロー1201)、最終的にノードAはParameter=0.5に決定している(フロー1202)様子を示している。ここで、ノードA、Bのデフォルト値が相違しているのは、例えばノードBはノードA以外にノードC、D、Eなどと接続されており、例えばノードC、Dの方はリンクの特性から或る程度Conservativeに行いたいという事情をネットワーク管理者が考慮したことによる。なお、各ノードに対して上述のように1つのデフォルト値を設定する以外に、Configurationの量は増えるが、各ノードに対して隣接相手ノード毎にデフォルト値を設定することも可能である。

【0034】リンク情報要約手段221は、上述のようにして得られたAggregationパラメータ情報と、図4に例示したような隣接ノード間リンク情報211と、更にリソース管理部203で管理されている情報とから、複数の物理リンクを要約し、そのリンク属性とし

て新規のQOS値の計算を行い、Aggregateリンク情報等を、自ノード用リンクステータデータベース202に格納する。

【0035】自ノードリンクステータデータベース202に格納されるデータの例を図6に示す。この例では、図4のNode ID-Aとの間の物理リンク3本が1本の論理リンク（論理My port ID-1 - 論理Remote port ID-1）に要約され、図4のNode ID-Bとの間の物理リンク3本が1本の論理リンク（論理My port ID-2 - 論理Remote port ID-1）に要約されている。そしてそれぞれの論理リンクに対して、Aggregationの計算の結果得られるトラフィッククラス、QOS情報の値を図6のように格納している。ここで格納しているQOS-A-1、QOS-B-1などが、Aggregationパラメータ $(0 \leq r \leq 1)$ によって、Aggressive、ConservativeにAggregateされたものである。これらはリソース管理部203で管理されている各物理リンクの帯域、遅延などのリソース使用量、空き容量の変化に応じてリンク情報要約手段221によって適宜更新される。

【0036】リンクステータルーティングプロトコル手段200は、自ノード用リンクステータデータベース202中のAggregateされたリンク情報に大きな変化があった場合に他のノードとトポロジー情報の交換を行う。ここで、大きな変化か否かを定める基準があまり小さすぎると、感度が良すぎるためにちょっと帯域情報が変化したことに基づくAggregation情報もFlooding（Broadcast）されることになり、反対に、あまり大きすぎると感度が悪くなり、ネットワークの状況が変わってもなかなか他のノードに通知できなくなってしまう。そこで、その基準はパラメータによってネットワーク管理者が変更できるようにしておくことが望ましい。例えば、もともと図6のようにもっていたAggregateされた結果のQOS-A-1などの値が、前回の値の50%変化したら、つまり、 $(QOS-A-1)/2$ 以下、 $(QOS-A-1)/3/2$ 以上になったら、大きな変化があったと見なすなどの処理が可能である。なお、リンクステータルーティングプロトコル手段200は、定期的に他のノードとトポロジー情報を交換するように構成することもできる。

【0037】また、リンクステータルーティングプロトコル手段200は、他のノードから受信したトポロジー情報を他ノード用リンクステータデータベース201に格納する。リンクステータルーティングプロトコル手段200は、自ノード用リンクステータデータベース202、他ノード用リンクステータデータベース201の両者を用いてネットワーク全体のトポロジーを把握し、それに基づき経路計算を行う。つまり、全てのノードは、

自リンクステータ情報を計算し、それをお互いに交換することで、全てのノードが全ネットワークの接続状態ならびに帯域情報などを知ることができ、これをベースに経路計算を行うことができる。

【0038】次にリンク情報Aggregation手段221におけるリンク情報の要約処理について詳しく説明する。

【0039】図7は、リンク情報Aggregation手段221が、複数の物理リンクのQOSをどのように1本の論理リンクのQOSとして表現するのかの例を示す。図7の上半分は、ATMノード300、301、302、303から構成されるネットワークにおいて、特にノードA、4 303のトポロジー情報交換装置が複数の物理リンクをAggregateしなかった場合の隣接リンク情報321と、ネットワーク内トポロジー情報320とを示し、図7の下半分はAggregateした場合の同様の情報323、322を示している。ここでは、隣接リンク情報321、323と、ネットワーク内トポロジー情報320、322とが、各リンクの属性として帯域を持つ場合を例に示す。隣接リンク情報321、323は、図2の隣接ノード間リンク情報211とリソース管理部203中の情報とを合体した情報に相当する。また、ネットワーク内トポロジー情報320、322は、図2の他ノード用リンクステータデータベース201の内容と自ノード用リンクステータデータベース202の内容とを合体したものに相当する。

【0040】ノードA、4 303とノードA、3 302との間の物理リンク310は3本存在し、それぞれ100Mbps、50Mbps、150Mbpsの帯域を持っている。もし、複数の物理リンクをAggregateせずに、それぞれのリンク情報を他のノードA、1、A、2、A、3へ転送した場合、全ノードがネットワーク内トポロジー情報320と同じ情報を持つことになる。

【0041】一方、複数の物理リンク310を1本の論理リンクにAggregateする場合を、ATMノード304、305、306、307から構成されるネットワークで示す。ここでは、ノードA、4とノードA、3の間の複数の物理リンクを1本の論理リンクにAggregateする場合を示し、リンク311のような形に認識されている。その論理リンクに対する新規のリンク属性は、隣接リンク情報323に示すように、100Mbps、50Mbps、150Mbpsの帯域の中の最大値として150Mbpsのように計算で求められる。しかしながら、リンク属性は1つではなく、複数存在するのが一般的であり、このような簡単な計算で求められるとは限らない。それについて以下に示す。

【0042】図8は、リンク属性について示したものであり、トラフィッククラスは、Constant Bit Rate (CBR)、Real Time Var

iable Bit Rate (rt-VBR), Non Real Time VBR, Available Bit Rate (ABR), Unspecified Bit Rate (UBR)などのクラスから構成され、ユーザが規定するQOSのmetricsは、Maximum Cell Rate (MaxCR), Available Cell Rate (AvCR), Cell Delay Variation (CDV), Cell Transfer Delay (CTD), Cell Loss Ratio (CLR)などの種類が存在し、ネットワークが規定するリンクコストとしてAdministrative Weight (AW)などがある。

【0043】各物理リンクでは、これらの複数の属性を持つ（全部持つとは限らない）ために、論理リンクにAggregateする場合、大きく分けて2段階の手順をとる。図9は、1段階目の手順であり、複数の物理リンクの属性をトラフィッククラス毎に整理する段階である。図9は、リンク1, 2, 3の3本の物理リンクが、トラフィッククラスA, B, C, D (A~Dはそれぞれ異なるトラフィッククラスを意味し、実際には図8に例示したようなトラフィッククラスになるが、ここでは便宜上、A~Dで表記している)のうち、一部をそれぞれサポートし、且つそれぞれのトラフィッククラスにおいてサポートするQOSが、QOS-1, QOS-2, ..., QOS-8のようにさまざまな範囲に分かれていることを想定している。図9に示すように、この3本の物理リンクを1本の論理リンクにAggregateする場合、その論理リンクはトラフィッククラスA, B, C, Dの全てをサポートするものとする。そして、例えば、トラフィッククラスAに着目すると、Link-1はQOS-1, Link-2はQOS-3, Link-3はQOS-6をそれぞれ持つため、トラフィッククラスAは、QOS-1, QOS-3, QOS-6からAggregation値を計算すれば良いことが判る。つまり、この段階は、Aggregateすべき、QOSの候補を抽出する段階である。他のトラフィッククラスについても同様にQOSの計算すべき候補を抽出することができる。

【0044】図10は、2段階目の手順であり、或るトラフィッククラス（図9と異なり、図10では具体的なトラフィッククラスを例示している）に着目し、そのトラフィッククラスにおけるAggregateすべきQOSの候補が、QOS1, QOS2, QOS3と例えば3つある時に、それらをAggregateする場合に、AggressiveなAggregation（過大評価）とConservativeなAggregation（過小評価）とを示している。QOS1, QOS2, QOS3はそれぞれ、MaxCR, AvCR, CDV, CTD, CLR, AWらの複数の属性から

構成されており、それらの値が図示されている。これらのリンクの属性には、大きく分けて、自ノードから宛先ノードまでmetricsをパスに沿って足し合わせていくことによって最適経路計算するAdditive Metricsと、各リンクの値そのものの大きさのみに意味を持つNon-Additive Metricsとがある。Additive Metricsは、CDV, CTD, CLR, AWで、Non-Additive Metricsは、MaxCR, AvCRである。Aggressive Aggregationの場合、Non-Additive Metricsは複数のリンクのうちから最大値をとり、Additive Metricsは最小値をとるようにすればよい。逆に、Conservative Aggregationの場合、Non-Additive Metricsは複数のリンクのうちから最小値をとり、Additive Metricsは最大値をとるようにすれば良い。

【0045】しかしながら、現実には、この両極端ではなく、その中間の値が使われることが一般的であり、それを表すのが、図2のAggregationパラメータ交換手段220によって交換されるパラメータである。このパラメータrは、例えば0から1の実数をとらせ、r=0のときはConservative、r=1のときはAggressive、その中間的な値rのときは、Aggressiveな値XとConservativeな値Yとから線形補間して $r \times X + (1-r) \times Y$ のように値を求めることにより、このような中間的な値を求めることができる（図10中に図示）。

【0046】図11は第2の発明にかかるトポロジー情報交換装置の実施の形態の機能構成を示し、図2と同一符号は同一部分を示し、700は性能評価手段、701はAggregationパラメータ計算手段である。本実施の形態のトポロジー情報交換装置は、図2に示した実施の形態のトポロジー情報交換装置の構成に、更に2つの機能手段700, 701が追加されている。なお、図11の装置は、例えばCPU、メモリ等を有するコンピュータと制御プログラムを記録した記録媒体（磁気ディスクや半導体メモリ等）とで実現され、記録媒体中の制御プログラムがコンピュータに読み込まれ、コンピュータの動作を制御し、コンピュータを図11に示す10個の機能手段として機能させる。

【0047】性能評価手段700は、或る時点のAggregationパラメータによってリンク情報がAggregateされている時に、そのAggregateリンクにおける呼損率、Aggregateリンクネットワーク利用率を評価する。このような評価は常時実行され、定期的に評価結果がAggregationパラメータ計算手段701に与えられる。

【0048】ここで、それぞれ自ノードは自分の隣接ノードに関するAggregateリンクの論理ポート番

号情報と、実際の物理ポート情報との対応を図6のように管理しているため、性能評価手段700はその情報を参照することでAggregateリンクが何であるかを認識することができる。

【0049】また、Aggregateリンクにおける呼損率とは、Aggregateリンクを利用してコネクションを設立しようとした全コネクション数のうち、リソース不足等のためコネクション設定に失敗したコネクションの確率のことである。この呼損率は、そのAggregateリンクを通過しようとするコネクション要求の数と、失敗したコネクションの数とを累計していくことで簡単に求めることができる。図12に示すように、Aggregationパラメータを1に近づけ、あまりAggressiveにリソース情報を周りにFloodingすると、このリンクを通過するトラフィックが増えるため、コネクション設定が失敗し易くなり、呼損率が大きくなる傾向がある。

【0050】更に、Aggregateリンクネットワーク利用率とは、Aggregateされた物理リンク全ての容量に対して、単位時間当たりどれくらいの帯域を使用しているかを示す確率のことであり、利用率が1（最大）ということは完全に使い切っていることを意味し、逆に0（最小）であれば全然有効利用されていないことを意味する。図12に示すように、Aggregationパラメータを0に近づけ、あまりConservativeにやりすぎると、おそらく利用率は0に近づき、逆にAggressiveにやりすぎると、利用率は1に近づき、前述したように非常に呼損率の大きい状態になってしまう。

【0051】従って、適当な値のAggregationパラメータであれば、利用率が十分に高く然も呼損率が小さい状態に保つことが可能となる。Aggregationパラメータ計算手段701は、性能評価手段700によって求められた呼損率、Aggregateリンクネットワーク利用率の値に基づき、呼損率をなるべく下げ、またAggregateリンクネットワーク利用率をなるべく高める方向に、Aggregationパラメータを変更する。具体的には、性能評価手段700から新たな評価結果が送られてくる毎に、呼損率およびAggregateリンクネットワーク利用率が所定の基準を満たしているかどうかを判定し、呼損率が基準を満たしていない場合には、Aggregationパラメータの値を所定値（例えば1/10）だけConservativeの方向に変更し、逆にAggregateリンクネットワーク利用率が基準を満たしていない場合には、Aggregationパラメータの値を所定値（例えば1/10）だけAggressiveの方向に変更することにより、Aggregationパラメータの値を最適値に収束させていく。

【0052】なお、リンクステートルーティングプロト

コル手段200が、定期的な交換でなく、自ノード用リンクステートデータベース202中のAggregateされたリンク情報に大きな変化があった場合に他のノードとトポロジー情報の交換を行う構成の場合、性能評価手段700の評価項目にAggregateリンクの経路情報量を追加し、Aggregationパラメータ計算手段701は、この経路情報量をも考慮してAggregationパラメータを計算するように構成することもできる。ここで、Aggregateリンクの経路情報量とは、単位時間当たりFloodingされて、受信したリンクステートの情報量のことであり、Floodingで受信したバケットのうち、Aggregateリンクの情報を記述しているもののみを累計することで求めることができる。ConservativeにAggregateしている場合とAggressiveにAggregateしている場合とで、どのくらい頻繁にAggregation情報が変わるかが決まり、図12に示すように、一般にAggregationパラメータが1でAggressiveな場合、経路情報量が増える傾向を示す。従って、Aggregationパラメータ計算手段701は、性能評価手段700から新たな経路情報量を送られてくる毎に、それが所定の基準を満たしているかどうかを判定し、基準を満たしていない場合には、Aggregationパラメータの値を所定値（例えば1/10）だけConservativeの方向に変更することにより、経路情報量をできるだけ少なくするような最適値にAggregationパラメータの値を収束させていく。

【0053】さて、Aggregationパラメータ計算手段701で計算された新規パラメータはAggregationパラメータ交換手段220へ渡され、Aggregationパラメータ交換手段220は、新規Aggregationパラメータを隣接ノード間で交換する。図13は、図11のAggregationパラメータ交換手段220が隣接ノード間で行うパラメータのNegotiationの方法の例について示す。ノードA、B間がすでにParameter=0.5で合意している時に（フロー1310、1311）、ネットワークの性能評価結果に伴い、ノードA側がParameterを変更してParameter=0.2にすると（フロー1312）、ノードBはParameter=0.2に合意している（フロー1313）。

【0054】上述のように隣接ノード間で交換されたAggregationパラメータはリンク情報Aggregation手段221に送られる。これにより、再び新規パラメータをベースにして、リンク情報Aggregation手段221が、Aggregationの値を変更し、自ノード用リンクステートデータベース202を更新する。そして、その内容はリンクステートルーティングプロトコル手段200における定期処理ま

たは内容に大きな変化があったときに、全てのノードに送出される。その後も性能評価手段700はネットワーク内の性能評価を行っており、上述した最適化処理が反復的に行われる。

【0055】図14は第3の発明にかかるトポロジー情報交換装置の実施の形態の機能構成を示し、図2と同一符号は同一部分を示し、900はAggregationグループ通知手段、901はリンクグループ作成手段、902はリンク情報Aggregation手段、903はAggregationグループ計算手段である。本実施の形態のトポロジー情報交換装置は、図2に示した第1の実施の形態のトポロジー情報交換装置の構成に、更に3つの機能手段900、901、903が追加されている。なお、この装置は、例えばCPU、メモリ等を有するコンピュータと制御プログラムを記録した記録媒体（磁気ディスクや半導体メモリ等）とで実現され、記録媒体中の制御プログラムがコンピュータに読み込まれ、コンピュータの動作を制御し、コンピュータを

図14に示す11個の機能手段として機能させる。

【0056】Aggregationグループ計算手段903は、複数の物理リンクの中からリンク属性の類似しているものをグループ化する機能をもつ。例えば、図15に示すように、ATMノード1000、1001間の物理リンクをリンク属性の類似度に応じてAggregation group1、2のように2つのグループに分ける。この結果、1本の論理リンクにAggregateするのではなく、ATMノード1000、1001間では2つの論理リンクにAggregateされることになる。ここで、本実施例においては、ATMノード1000、1001間の物理リンクをグループ化する権限は、ATMノード1000、1001に割り当てられたNode IDの大きな方のノードに与え、他方のノードはそれを受け入れるだけである。例えばATMノード1000の方がATMノード1001よりNode IDが大きい場合、ATMノード1000内のAggregationグループ計算手段903がグループ化する。また、ATMノード1001にATMノード1000とは別のATMノード（1002。図示せず）が接続されており、ATMノード1001のNode IDがATMノード1002より大きいとすると、ATMノード1001とATMノード1002間の物理リンクのグループ化はATMノード1001内のAggregationグループ計算手段903が行う。

【0057】具体的には、Aggregationグループ計算手段903は、図4に示したような隣接ノード間リンク情報211を参照して、Node IDが自ノードの方が大きい隣接ノードとの間の物理リンクを認識し、リソース管理部203で管理されている各物理リンクのリンク属性に基づき、リンク属性の類似しているもの同士をグループ化する。リンク属性が類似する物理リ

リンクとは、例えば、サポートするトラフィッククラスが同じで且つそれぞれのトラフィッククラスにおいてサポートするQOSが類似しているものや、サポートするトラフィッククラスがほぼ同じで且つそれらでサポートするQOSが類似しているもの等である。

【0058】Aggregationグループ計算手段903は、各グループに対して識別番号を付与し、或る隣接ノード間の物理リンクに関し、どのリンクグループがどのリンクとどのリンクかを示す例えば図16に示すようなデータを生成し、リンクグループ作成手段901およびAggregationグループ通知手段900に与える。

【0059】Aggregationグループ通知手段900は、このデータを該当する隣接ノードへ通知する。当該隣接ノードのAggregationグループ通知手段900は、受信したデータを自ノードのリンクグループ作成手段901に通知する。

【0060】リンクグループ作成手段901は、図16に示したようなデータを自ノードのAggregationグループ計算手段903およびAggregationグループ通知手段900を通じて他ノードのAggregationグループ計算手段903から受け取ると、そのデータに従い、隣接ノード間リンク情報211とリソース管理部203の内容とに基づいて実際にリンクグループを作成し、リンク情報Aggregation手段902に通知する。リンク情報Aggregation手段902は、作成されたリンクグループ毎に、Aggregationパラメータ交換手段220で隣接ノード間で交換されたAggregationパラメータを基準にAggregationを行い、自ノード用リンクステートデータベース202に登録する。

【0061】自ノード用リンクステートデータベース202に格納されるデータの例を図17に示す。図6とはほぼ同じ内容であるが、図6では図4のNode ID-Bとの間の物理リンク3本が1本の論理リンク（論理My port ID-2 - 論理Remote port ID-1）に要約されているのに対し、図17では2種類の論理リンク（論理My port ID-2 - 論理Remote port ID-1）と（論理My port ID-3 - 論理Remote port ID-2）とに要約されている点が相違している。これは、1本に要約すると余りにもトラフィック特性の偏りがあるため、2つのリンクグループに分割されたことによる。

【0062】なお、Aggregationパラメータ交換手段220が交換するAggregationパラメータとして、ノードの各リンクグループ毎に1つ持つようにし、各リンクグループ毎にAggregationパラメータをノード間で交換するようにしても良い。この場合、リンク情報Aggregation手段90

2はリンクグループ毎のAggregationパラメータを使って要約リンク情報を計算し直す。

【0063】図18は第4および第5の発明にかかるトポロジー情報交換装置の実施の形態の機能構成を示し、図11および図14と同一符号は同一部分を示す。本実施の形態のトポロジー情報交換装置は、図14に示した実施の形態のトポロジー情報交換装置の構成に、更に2つの機能手段700、701が追加されている。なお、この装置は、例えばCPU、メモリ等を有するコンピュータと制御プログラムを記録した記録媒体（磁気ディスクや半導体メモリ等）とで実現され、記録媒体中の制御プログラムがコンピュータに読み込まれ、コンピュータの動作を制御し、コンピュータを図18に示す13個の機能手段として機能させる。

【0064】性能評価手段700は、或る時点のAggregationパラメータによってリンク情報がAggregatedされている時に、そのAggregateリンクにおける呼損率、Aggregateリンクネットワーク利用率、及び必要に応じて更に経路情報量を評価する手段である。基本的な動作は図11で説明したものと同じであるが、本実施の形態における性能評価手段700が評価対象とするAggregateリンクは、リンクグループ作成手段901でリンクグループ化された個々のAggregateリンクである。リンクグループ毎のAggregateリンクが何であるかは、図4に示したような隣接ノード間リンク情報211とリンクグループ作成手段901に与えられる図16に示すようなデータとを統合した図19に示すようなデータを管理しておくことで認識できる。また、Aggregateリンクにおける呼損率、Aggregateリンクネットワーク利用率、経路情報量は、それぞれの物理リンク毎の呼損率X、Aggregateリンクネットワーク利用率Y、経路情報量Zを計算し、そこで得られたX、Y、ZをAggregateリンク全体に渡って、再計算することで求めることができる。性能評価手段700で定期的に求められた評価結果は、Aggregationパラメータ計算手段701およびAggregationグループ計算手段903に与えられる。

【0065】Aggregationパラメータ計算手段701は、性能評価手段700によって求められた呼損率、Aggregateリンクネットワーク利用率、経路情報量の値に基づき、呼損率や経路情報量をなるべく下げ、またAggregateリンクネットワーク利用率をなるべく高める方向に、Aggregationパラメータを変更する。その動作は第2の発明の実施の形態の場合と基本的に同じであるが、本実施の形態の場合、Aggregationパラメータは各リンクグループ毎に1つあり、Aggregationパラメータ計算手段701は各リンクグループ毎のAggregationパラメータの値を求める点が相違する。また、

Aggregationパラメータ交換手段220は、Aggregationパラメータ計算手段701で計算された新規パラメータを隣接ノード間で交換してNegotiationを行い、結果をリンク情報Aggregation手段902に通知する。

【0066】他方、Aggregationグループ計算手段903は、性能評価手段700の評価結果により、呼損率や経路情報量をなるべく下げ、かつAggregateリンクネットワーク利用率をなるべく高めるように、リンクグループを変更する。具体的には、性能評価の結果、複数のリンクグループのうち所定の基準よりも性能が劣化しているグループがあるかどうかを調べる。若し、性能が劣化しているリンクグループがあれば、そのリンクグループ中の物理リンクのQOS特性を調べ、分離した方が性能が上がるかどうかを判断し、性能があがると判断されたら分離し、そうでなければ分離せずにそのままにする。このような判断を定期的に行うことにより、徐々に最適なリンクグループへと収束させる。なお、Aggregationグループ計算手段903が再グループ化を行うのは、Node IDが自ノードの方が大きい隣接ノードとの間の物理リンクである。

【0067】Aggregationグループ計算手段903は、変更後の各グループに対して識別番号を付与し、前述した図16に示すようなデータを生成し、リンクグループ作成手段901およびAggregationグループ通知手段900に与える。Aggregationグループ通知手段900は、このデータを該当する隣接ノードへ通知し、当該隣接ノードのAggregationグループ通知手段900は、受信したデータを自ノードのリンクグループ作成手段901に通知する。

【0068】リンクグループ作成手段901は、図16に示したようなデータを自ノードのAggregationグループ計算手段903およびAggregationグループ通知手段900を通じて他ノードのAggregationグループ計算手段903から受け取ると、そのデータに従い、隣接ノード間リンク情報211とリソース管理部203の内容とに基づいて実際にリンクグループを作成し直し、リンク情報Aggregation手段902に通知する。リンク情報Aggregation手段902は、作成されたリンクグループ毎に、Aggregationパラメータ交換手段220で隣接ノード間で交換された各リンクグループ毎のAggregationパラメータを基準にAggregationを行い、自ノード用リンクステートデータベース202に登録する。

【0069】以上の各実施の形態は、一般に、同一レイヤ内でのリンクのAggregationに対し適用することができる。以下では、下位レイヤのトポロジー情

報を上位レイヤに見せる場合のAggregationについて説明する。

【0070】図20は本発明のトポロジー情報交換装置を適用する階層的なPNNIルーティングのドメイン（ピアグループ；Peer Group）構成を示す。まず、同図を参照して、下位レイヤのトポロジー情報を上位レイヤに見せる仕組みの概要を説明する。

【0071】図20において、「物理網」と書かれている下の部分が実際の物理スイッチやリンクを表しており、Peer Group1400、1401、1402といったドメインから構成され、それぞれのPeer Groupは複数のノードとリンクから構成されている。他方、「論理網」と書かれている上の部分は、実際には論理的にそのような網があるかのようにルーティング情報を交換しようというための網を表している。物理網において、各Peer Groupに存在する代表ノード（Peer Group Leader）1480等が、お互いの間にVC（Virtual Circuit）と呼ぶコネクション1470、1471、1472を設定し、そのコネクション上で経路情報を交換することにより、上述の論理網があたかも構築されているように見える。

【0072】代表ノード間で1本ずつ張られたVC上では、それぞれのPeer Group間に存在しているリンク全てを認識し交換する。例えばPeer Group1400、1401の間には、1440、1441の2本の物理リンクが存在するが、その情報をVC1470上で交換する。この交換は、図21のように、最下位のレベルのHello packetに追加して、Peer Group間にどのようなリンクがあるのかを記述することで行われる。

【0073】図21において、My Node ID、Remote Node IDは、それぞれPeer Group LeaderのノードIDを意味し、それぞれのMy Node ID、Remote Node IDは、論理リンク用のHelloを意味するものとして、特別な値0xFFFFFFFFが入る。その基本的なHello要素の後ろに、Peer Group間の複数リンクを記述するフィールドがくる。最初にLogical Link数、例えばPeer Group1400、1401間には2つのリンク1440、1441があるので、この値は2となる。そして、My Border Node ID、Remote Border Node IDは、1440あるいは1441の境界に存在するBorder Nodeのそれぞれの最下位レベルのNode IDを意味する。またMy Border port ID、Remote Border IDも、1440、1441のリンクがそれぞれのBorder Nodeのどのポートから出ているかを記述している。これらBorder Node ID、Bo

der port IDの2つの組で一意にPeer Group間のリンクを識別できる。そしてそれらのリンクに対して、Peer Group Leaderは、ユニークな論理port IDを割り当て、それをMy logical port ID、Remote logical port IDの値として記述する。また、Logical link group IDは、論理リンクをAggregateするために利用するIDであり、もし同じIDであればそのリンクをAggregateすることになる。

【0074】上記のようなHelloメッセージをPeer Group Leader間のVC上で交換し、例えば図22に示すような格納形式で管理することにより、上位階層において、どのような論理リンクが論理ノード1410、1411間でどのような論理ポート番号において接続され、かつそれぞれのLogical link group IDは何なのかを判断することができる。また、Logical link group IDが同じ論理リンクはAggregateされて1本に縮退でき、その場合、また論理ポート番号等が再決定される。

【0075】さて、図20において、Peer Group間でトポロジー情報を交換する場合、Peer Group内の代表ノード1480等は、自Peer Group内のすべてのトポロジー情報を他のPeer Groupに対して開示するのではなく、ノードとリンクから構成されるトポロジー情報をAggregateして上位レイヤ（Peer Group）においてトポロジー情報の交換を行う。この例では、Peer GroupX、1（1400）は、上位Peer Groupにおいて、論理ノードX、1、Peer GroupX、2（1401）は、論理ノードX、2、Peer GroupX、3（1402）は論理ノードX、3として表現され、論理ノードX、1、X、2、X、3間が複数の論理リンクで接続されているような形でトポロジー情報が交換され、この情報はそれぞれの下位のPeer Group1400、1401、1402内のすべてのノードへと提供されている。

【0076】ここで論理ノードX、1、X、2の間の論理リンク1450、1451はそれぞれ物理リンク1440、1441に対応し、論理ノードX、1、X、3の間の論理リンク1452は、物理リンク1442に対応する。論理リンク1450、1451、1452は、Aggregateしていないリンクである。一方、論理ノードX、2、X、3間の論理リンク1453は、物理リンク1443、1444が一本にAggregateされていることを示している。

【0077】例えば、Peer GroupX、2（1401）のトポロジーを代表ノード1480がAggregateして表現する場合を例にあげると、他のPe

erGroupとの間の境界ノード(Border Node) 1420, 1421, 1422, 1423を選択し、それをStar型に接続して、最初に1490のような形に近似を行う。その後上位Peer Groupにおける論理リンクがAggregateされた結果できたかどうかを判断し、ここでは論理リンク1453がAggregateされた結果できているため、近似トポロジー上から、上記境界ノード1421, 1422もAggregateして近似トポロジーを求め直すと、1491のような形になる。

【0078】図23は第6および第7の発明にかかるトポロジー情報交換装置の実施の形態の機能構成を示す。本実施の形態のトポロジー情報交換装置は、図20で述べた手順を実現する機能を持ち、特に1480のようなPeer Group内の代表ノードが本機能を有する。本トポロジー情報交換装置は、リンクステートルーティングプロトコル手段1500、下位Peer Group用リンクステートデータベース1501、上位Peer Group用リンクステートデータベース1502、Helloプロトコル手段1510、下位Peer Group用隣接ノード間リンク情報1511、上位Peer Group用隣接ノード間リンク情報1512、下位Peer Group Border Node検出手段1520、中間形式作成手段1521、Topology Aggregation計算手段1522およびBorder Node Aggregation手段1523の10個の機能手段から構成される。なお、図23に示す装置は、例えばCPU、メモリ等を有するコンピュータと制御プログラムを記録した記録媒体(磁気ディスクや半導体メモリ等)とで実現され、記録媒体中の制御プログラムがコンピュータに読み込まれ、コンピュータの動作を制御し、コンピュータを図23に示す10個の機能手段として機能させる。

【0079】以下、Peer Group 1401内の代表ノード1480が本機能を有する場合を例にして、各機能を詳細に説明する。

【0080】ハロー(Hello)プロトコル手段1510は、図20におけるPeer Group 1401内の隣接ノードだけでなく、上位Peer Group 1404内の隣接ノードに対しても、図21で説明したような方法で制御メッセージのやりとりを行う。ハロー(Hello)プロトコル手段1510が隣接ノード間の複数の全ての物理リンクにおいて制御メッセージのやりとりを行い、下位Peer Group用隣接ノード間リンク情報を認識した場合、下位隣接ノード間リンク情報1511にその情報を格納する。この下位隣接ノード間リンク情報1511は第1乃至第5の発明の実施の形態における隣接ノード間リンク情報211に相当する。一方、ハロー(Hello)プロトコル手段1510が上位Peer Groupにおける隣接論理ノード

間の複数の全ての論理リンクにおいて制御メッセージのやりとりを行い、隣接論理ノード間に存在する論理リンクを認識した場合、例えば前述した図22に示したような形式で、上位Peer Group用隣接ノード間リンク情報1512にその情報を格納する。

【0081】リンクステートルーティングプロトコル手段1500は、下位Peer Group内の他のノードとトポロジー情報を交換することにより、下位Peer Group用リンクステートデータベース1501に、そのトポロジー情報を保存する。この下位Peer Group用リンクステートデータベース1501は第1乃至第5の発明の実施の形態における他ノード用リンクステートデータベース201および自ノード用リンクステートデータベース202に相当する。一方、下位Peer Group内の代表ノード(Peer Group Leader)は、リンクステートルーティングプロトコル手段1500により、上位Peer Group内の他の論理ノードとトポロジー情報を交換することが可能であり、それにより上位Peer Groupのリンクステートデータベース1502に、そのトポロジー情報を保存する。下位と上位のPeer Group内のリンクステートデータベースを総合して、宛先への経路を計算することが可能となる。

【0082】ここで、下位Peer Group用リンクステートデータベース1501を元にして、下位Peer Group Border Node検出手段1520は、他のPeer Groupと隣接しているBorder Nodeを検知し、その結果を中間形式作成手段1521に通知する。中間形式作成手段1521はこの通知に従い中間形式のトポロジー情報を作成する。以下、中間形式作成手段1521の処理を図24を用いて説明する。

【0083】図20のPeer Group 1401内の物理トポロジーを1601とし、その中で他のPeer Groupと接しているBorder Nodeを1, 2, 3, 4のノードとする。そのとき、1, 2, 3, 4のノード間をFullmeshにリンクで接続し、それぞれのリンクごとの帯域や遅延などのQOSを実際の物理トポロジーの性能に近似するように計算を行って割り当てる。この近似計算によって1602を得る処理が、中間形式作成手段1521の行う中間形式作成処理である。

【0084】具体的には、先ず図24の1601のトポロジーにおいて、Border Node 1, 2, 3, 4に注目し、Border Node間の最小コスト経路を求める。このとき若し、複数ある場合には複数の経路を求め、諸条件を考慮してその中の1つを選択する。例えばBorder Node 1, 2間のリンクコストc14を求める場合、図25に示すように1→4への経路のうち合計コストが同じ20となる最小コスト経路が

2本あったとする。選ばれたリンクコストが最小のこの2つのパスにおいて、今度は他のQOS情報、例えば帯域、あるいは遅延などのコストも計算し、いずれがリンクが多いかによって、2本のうちから1本の最小コスト経路を選択し、その経路におけるリンクコストを c_{14} とする。 c_{14} は複数のQOS、つまりセル廃棄率 (Cell Loss Ratio: CLR), セル転送遅延 (Cell Transfer Delay: CTD), セル遅延ゆらぎ (Cell Delay Variation: CDV), 最大リンク容量 (Maximum Cell Rate: MaxCR), 残余帯域 (Available Rate: AvCR) 等から構成される、いわばベクトル情報である。同様にして図24の中間形式におけるリンクコスト c_{12} , c_{13} , c_{23} , c_{24} , c_{34} を求めることができる。

【0085】次に、Topology Aggregation計算手段1522は、上記中間形式から実際のTopology表現形式に変換する。Topology表現形式には、例えば3通りが想定され、図24の1603のような半径のQOSの値が全て同一の値をとるスター型の形式 (Symmetric Star) か、1604のような半径のQOSの値が異なる値をとるスター型の形式 (Asymmetric Star) か、1605のようなAsymmetricStar型のトポロジーにさらにBorder Node間にバイパス経路を加えた形式 (Asymmetric Star+Exception) という3通りの近似方法があり、上記の中間形式から線形計画法などを用いて、理論的に近似計算を行う。下位Peer Groupのノードとリンクの接続形態 (トポロジー) をAggregateした結果の近似精度は、1603のSymmetric Star, 1604のAsymmetric Star, 1605のAsymmetric Star+Exceptionの順に高くなる。従って、近似精度の閾値を予め設定し、その閾値を満足している方式を選択する。Symmetric Star, Asymmetric Star, Asymmetric Star+Exceptionの順に計算量が複雑

になるため、Symmetric Starで処理ができれば一番よい。しかし、これでは近似精度の閾値を満足しなければ、次に複雑なAsymmetric Starで試みて、それでもだめならAsymmetric Star+Exceptionという形で段階的に選択する。

【0086】以下に中間形式のトポロジー表現形式から実際のTopology表現形式に変換する処理を説明する。

【0087】QOSには大きくわけてセル廃棄率 (CLR), セル転送遅延 (CTD), セル遅延ゆらぎ (CDV) のような経路上のすべてのリンクのパラメータを累計しその合計が要求品質を満足するかを判断するAdditive Parameterと、最大リンク容量 (MaxCR), 残余帯域 (AvCR) のように各リンクごとの品質が要求品質を満足するかどうかのみ判断すればよい、Non-Additive Parameterとの2種類がある。各リンクはこれらの複数のQOS Parameterからなるベクトルで表現され、このベクトルで記述されたリンクを図24の(c), (d), (e) のような形態で近似する。基本的にそれぞれのベクトル要素、QOSパラメータは独立に取扱

い、上記2種類のパラメータでそれぞれ処理方法が異なる。

【0088】まず式を導入するために必要な情報を定義する。 r_i ベクトルは、図24の(c), (d),

(e) の近似後の各リンクごとのLink metricsであり、 C_{ij} ベクトルは図24の中間形式(b)のLink metricsである。以下で、 r_i (MaxCR), C_{ij} (MaxCR) などと記述しているのは、ベクトルのうちのMaxCRの部分の要素値という意味である。つまり要素値はここではスカラーとなる。基本的には中間形式(b)と形態(d)との間の近似解を求めるのが基本となる。形態(c), (e)については、後述する。

【0089】中間形式(b)から形態(d)を求めるための基本式は次の通りである。

【数1】

$$\text{Min}\{\vec{r}_i(\text{MaxCR}), \vec{r}_j(\text{MaxCR})\} \leq \vec{c}_{ij}(\text{MaxCR})$$

$$\text{Min}\{\vec{r}_i(\text{AvCR}), \vec{r}_j(\text{AvCR})\} \leq \vec{c}_{ij}(\text{AvCR})$$

$$(\vec{r}_i(\text{CLR}) + \vec{r}_j(\text{CLR})) \geq \vec{c}_{ij}(\text{CLR})$$

$$(\vec{r}_i(\text{CDV}) + \vec{r}_j(\text{CDV})) \geq \vec{c}_{ij}(\text{CDV})$$

$$(\vec{r}_i(\text{CTD}) + \vec{r}_j(\text{CTD})) \geq \vec{c}_{ij}(\text{CTD})$$

$$\vec{r}_i \geq 0$$

$$(i, j=0, 1, 2, \dots)$$

【0090】式〔数1〕は、Aggregateをする場合に、帯域をなるべく小さく、遅延はなるべく大きく近似しようと試みており、情報を少し悪い方に求めようというアプローチであり、ConservativeにAggregateするのに相当する。一方、帯域はな*

＊なるべく大きく、遅延はなるべく小さく近似するアプローチも可能であり、その場合は次の式〔数2〕を計算することにより、Aggressive Aggregation値を求めることができる。

【数2】

$$\text{Min}\{\vec{r}_i(\text{MaxCR}), \vec{r}_j(\text{MaxCR})\} \geq \vec{c}_{ij}(\text{MaxCR})$$

$$\text{Min}\{\vec{r}_i(\text{AvCR}), \vec{r}_j(\text{AvCR})\} \geq \vec{c}_{ij}(\text{AvCR})$$

$$(\vec{r}_i(\text{CLR}) + \vec{r}_j(\text{CLR})) \leq \vec{c}_{ij}(\text{CLR})$$

$$(\vec{r}_i(\text{CDV}) + \vec{r}_j(\text{CDV})) \leq \vec{c}_{ij}(\text{CDV})$$

$$(\vec{r}_i(\text{CTD}) + \vec{r}_j(\text{CTD})) \leq \vec{c}_{ij}(\text{CTD})$$

$$\vec{r}_i \geq 0$$

$$(i, j=0, 1, 2, \dots)$$

【0091】以下では、Conservative Aggregationの場合を例にとって更に式を変

形、簡略化すると、次式〔数3〕が得られる。
【数3】

31

32

$$\vec{r}_i(\text{MaxCR}) = \text{Min}\{\vec{c}_{ij}(\text{MaxCR})\}, j=0, 1, 2, \dots$$

$$\vec{r}_i(\text{AvCR}) = \text{Min}\{\vec{c}_{ij}(\text{AvCR})\}, j=0, 1, 2, \dots$$

Min : x

subject-to

$$(\vec{r}_i(\text{CLR}) + \vec{r}_j(\text{CLR})) \geq \vec{c}_{ij}(\text{CLR}), i, j=0, 1, 2, \dots$$

$$(\vec{r}_i(\text{CLR}) + \vec{r}_j(\text{CLR})) - \vec{c}_{ij}(\text{CLR}) \geq x, i, j=0, 1, 2, \dots$$

Min : y

subject-to

$$(\vec{r}_i(\text{CTD}) + \vec{r}_j(\text{CTD})) \geq \vec{c}_{ij}(\text{CTD}), i, j=0, 1, 2, \dots$$

$$(\vec{r}_i(\text{CTD}) + \vec{r}_j(\text{CTD})) - \vec{c}_{ij}(\text{CTD}) \geq y, i, j=0, 1, 2, \dots$$

Min : z

subject-to

$$(\vec{r}_i(\text{CLR}) + \vec{r}_j(\text{CLR})) \geq \vec{c}_{ij}(\text{CLR}), i, j=0, 1, 2, \dots$$

$$(\vec{r}_i(\text{CLR}) + \vec{r}_j(\text{CLR})) - \vec{c}_{ij}(\text{CLR}) \geq z, i, j=0, 1, 2, \dots$$

$$\vec{r}_i \geq 0$$

【0092】以上より、最大リンク容量 (MaxCR)、残余帯域 (AvCR) は、 r_i の値を求めるときには、 C_{ij} (i 固定、 $j=1, 2, \dots$) の最小値を求めることで簡単に求められる。一方、セル廃棄率 (CLR)、セル転送遅延 (CTD)、セル遅延ゆらぎ (CDV) 等は、上記のような線形計画法に変換することにより解くことができる。

【0093】さて、ここで、形態 (c) は形態 (d) の特別な場合であり、 $r_i = r_j$ という場合に簡略化されるので、上記の式〔数3〕の右の項は容易に変形して簡略化される。

【0094】形態 (c) は、形態 (d) で求めた値を使って逆に中間形式の C_{ij} を求め直すと、中間形式の元の C_{ij} の値とは多少異なる場合がある。そのずれの大きさが相対的にもっとも大きいリンクを選び、そのリンクの近似精度を上げる試みを行う。つまり、そのリンクが $k-1$ とすると、形態 (e) において、リンク $k-1$ に図24 (e) のように、バイパスリンクを作成し、その新たなリンクにおいて、再び計算式を定義して、計算することとなる。計算式は (d) で求めた式と同様に容易に作成して求めることができる。これによって、求められた r_i においてまた、中間形式の C_{ij} を求め直

し、もとの中間形式とのずれが或る閾値以下、あるいはバイパスリンク数が或る閾値を越えると、収束したと見なし、処理を終了する。

【0095】本実施の形態では、図24 (a) から同図 (b) の中間形式への変換と、(b) から最終形態 (c)、(d)、(e) への変換との2段階の手順を踏んでいる。これは、形態 (c)、(d)、(e) への変換は線形計画法に基づいて計算されるが、あまりに複雑なトポロジーだとしらみつふして全ての経路を総当たり的に調べる線形計画法では、計算量が膨大に増えてしまうため、前準備として (b) の中間形式に変換し、

(b) → (c)、(d)、(e) の部分について必要となる線形計画法による計算量を削減するためである。

【0096】さて、Topology Aggregation 計算手段1522で図24の (c)、(d)、(e) の結果が得られた後の処理としては、2通り考えられる。その1つは、第6の発明に関するもので、得られた結果を、そのまま上位 Peer Group 用リンクステートデータベース1502中の、自論理ノードの内部トポロジー表現データとして保存することである。

【0097】他の1つは、第7の発明に関するもので、さらに Topology Aggregation 計算

30

40

50

手段1522で得られた結果を、Border Node Aggregation手段1523に渡して以下の処理を行わせることである。

【0098】Border Node Aggregation手段1523は、上位Peer Group用隣接ノード間リンク情報1512から、複数のBorder Node間のリンクのいずれがリンクAggregationしているかを調べ、もしAggregationしている場合には、上記Aggregation表現で用いたBorder Nodeのうち、該当するBorder NodeのAggregationを行う。例えば図26に示すように、X、2の下位トポロジーのTopology Aggregationが通常の処理で進むと、Port1、2、3、4からなるトポロジーで表現が可能となる。各Port1、2、3、4は実は下位のX、2のBorder Node1、2、3、4に相当している。ここで、X、2、X、3の間には2本のリンクが存在しており、それはPort2、3に接続するリンクである。もしこの2つのリンクのトラフィック特性が類似している場合、リンクのAggregationを行う。このとき、Border Node Aggregation手段1523は、Port2、3を1つのPortにAggregateする。

【0099】Border NodeのAggregationを行った場合、Border Node Aggregation手段1523は、再びQOSの値を計算し直す。即ち、Topology Aggregation計算手段1522で得られた要約トポロジーはPort1、2、3、4での近似を繰り返して得たものであり、Border Node2、3がAggregateされてPort数が1、2、4の3つに減ったため、今まで近似で各リンクごとに計算していたQOS値が変わってしまう。そこで、もう一度、前述した計算式を求め直し、各リンクのQOSを割り当て直す作業を行うことで、要約トポロジー情報を作成し直す。こうして得られた結果は、上位Peer Groupリンクステートデータベース1502中の、自論理ノードの内部トポロジー表現データとして保存される。

【0100】図27は第8の発明にかかるトポロジー情報交換装置の実施の形態の機能構成を示し、図23と同一符号は同一部分を示し、1700はAggregationパラメータ交換手段である。本実施の形態のトポロジー情報交換装置は、図23のトポロジー情報交換装置の構成に、さらに1つの機能手段1700が追加されている。なお、図27に示す装置は、例えばCPU、メモリ等を有するコンピュータと制御プログラムを記録した記録媒体（磁気ディスクや半導体メモリ等）とで実現され、記録媒体中の制御プログラムがコンピュータに読み込まれ、コンピュータの動作を制御し、コンピュータを図27に示す11個の機能手段として機能させる。

【0101】Aggregationパラメータ交換手段1700は、隣接論理ノードとAggregationパラメータを交換する。交換するAggregationパラメータの例を図28に示す。前述の〔数1〕、〔数2〕等の計算式で述べたように、Aggregationには大きく分けて、AggressiveとConservativeとがあり、それぞれの値を求めるための式を導出することができるため、何れで行うかを示すAggregation methodパラメータを交換する。さらに、それぞれの式において収束条件は、バイパスリンク最大数や近似精度閾値等の幾つかのパラメータによって制御されるため、それらを交換する。なお、Conservative AggregationとAggressive Aggregationとの間のどのくらいのAggregation割合を行うかについての値を交換することも可能である。

【0102】Aggregationパラメータ交換手段1700で交換された値は、Topology Aggregation計算手段1522に渡される。当該計算手段1522は、その値により、Aggregationの度合を調整して、計算し直す。計算し直されて得られたトポロジー情報は上位Peer Group用リンクステートデータベース1502中の、自論理ノードの内部トポロジー表現データとして保存される。

【0103】図29は第9の発明にかかるトポロジー情報交換装置の実施の形態の機能構成を示し、図27と同一符号は同一部分を示し、1800は性能評価手段、1801はAggregationパラメータ計算手段である。本実施の形態のトポロジー情報交換装置は、図27のトポロジー情報交換装置の構成に、さらに2つの機能手段1800、1801が追加されている。なお、図29に示す装置は、例えばCPU、メモリ等を有するコンピュータと制御プログラムを記録した記録媒体（磁気ディスクや半導体メモリ等）とで実現され、記録媒体中の制御プログラムがコンピュータに読み込まれ、コンピュータの動作を制御し、コンピュータを図29に示す13個の機能手段として機能させる。

【0104】性能評価手段1800は、ある時点のAggregationパラメータによってPeer Group内のトポロジー情報がAggregateされている時に、そのPeer Group内における呼損率、Peer Groupのネットワーク利用率を評価する。このような評価は常時実行され、定期的に評価結果がAggregationパラメータ計算手段1801に与えられる。

【0105】ここで、図30に示したように、各論理ノード毎のAggregateトポロジーは、各Portと中心点coreとの間のQOSパラメータ、並びに各Port間のパラメータ（Bypass exceptionsが存在する場合）によって記述できるため、こ

のようなデータを管理することにより、性能評価手段1800はAggregateトポロジーが何であるかを認識することができる。

【0106】また、Peer Group内における呼損率とは、Aggregateトポロジーを利用してそのPeer Groupを通じてコネクションを設立しようとした全コネクション数のうち、リソース不足等のためコネクション設定に失敗したコネクションの確率のことである。Peer Group内のトポロジー情報をあまりAggressiveにAggregateすると、そのPeer Groupを利用するトラフィックが増えるため、コネクション設定が失敗し易くなり、呼損率が大きくなる傾向がある。

【0107】更に、Peer Groupのネットワーク利用率とは、Peer Group内の物理リンク全ての容量に対して、単位時間当たりどれくらいの帯域を使用しているかを示す確率のことであり、利用率が1ということは完全に使い切っていることを意味し、逆に0であれば全然有効利用されていないことを意味する。Peer Group内のトポロジー情報をあまりConservativeにAggregateしすぎると、おそらく利用率は0に近づき、逆にAggressiveにやりすぎると、利用率は1に近づくが、前述したように非常に呼損率の大きい状態になってしまう。

【0108】なお、Peer Group内における呼損率およびPeer Groupのネットワーク利用率は、各ノードが分散的に累計をとることで求めることができる。但し、トポロジーがAsymmetric Aggregationなどの形態でAggregateされているため、どのAggregateトポロジーがどの物理リンクに対応し、どのような問題になっているかの判定が困難であるため、例えば、Aggregationに対応しているネットワーク全体のノードにおける呼損率、利用率を各Peer Group Leaderが周りのノードから集め、それに応じてネットワーク全体の統計情報を求めるようにする。

【0109】適当な値のAggregationパラメータであれば、利用率は1で然も呼損率が小さい状態に保つことが可能となる。Aggregationパラメータ計算手段1801は、性能評価手段1800によって求められた呼損率、ネットワーク利用率の値に基づき、呼損率をなるべく下げ、またネットワーク利用率をなるべく高める方向に、Aggregationパラメータを変更する。具体的には、性能評価手段1800から新たな評価結果が送られてくる毎に、呼損率およびAggregateリンクネットワーク利用率が所定の基準を満たしているかどうかを判定し、呼損率が基準を満たしていない場合には、Conservative Aggregation側に制御を移したり、閾値を高めて近似精度を上げたりし、逆にネットワーク利用率が基

準を満たしていない場合には、Aggressive Aggregation側に制御を移したりすることで、Aggregationパラメータの値を最適値に収束させていく。

【0110】なお、リンクステートルーティングプロトコル手段1500が、定期的な交換でなく、上位Peer Group用リンクステートデータベース1502中の、自論理ノードの内部トポロジー表現データに大きな変化があった場合に他の論理ノードとトポロジー情報の交換を行う構成の場合、性能評価手段1800の評価項目にAggregateトポロジーの経路情報量を追加し、Aggregationパラメータ計算手段1801は、この経路情報量をも考慮してAggregationパラメータを計算するように構成することもできる。ここで、Aggregateトポロジーの経路情報量とは、単位時間当たりFloodingされて、受信したリンクステートの情報量のことであり、Floodingで受信したパケットのうち、Aggregateトポロジーの情報を記述しているもののみを累計することで求めることができる。ConservativeにAggregateしている場合とAggressiveにAggregateしている場合とで、どのくらい頻繁にAggregation情報が変わるかが決まり、一般にAggressiveな場合、経路情報量が増える傾向を示す。従って、Aggregationパラメータ計算手段1801は、性能評価手段1800から新たな経路情報量が送られてくる毎に、それが所定の基準を満たしているかどうかを判定し、基準を満たしていない場合には、Conservative Aggregation側に変更することにより、経路情報量をできるだけ少なくするような最適値にAggregationパラメータの値を収束させていく。

【0111】さて、Aggregationパラメータ計算手段1801で計算された新規パラメータはAggregationパラメータ交換手段1700へ渡され、Aggregationパラメータ交換手段1700は、新規Aggregationパラメータを隣接論理ノード間で交換し、Conservative AggregationとAggressive Aggregationとの間でどのくらいのAggregation割合を行うかや近似精度閾値等についての値を交換する。この値はTopology Aggregation計算手段1522に送られる。これにより、再び新規パラメータをベースにして、Topology Aggregation計算手段1522が、Aggregationの値を変更し、上位Peer Group用リンクステートデータベース中の、自論理ノードの内部トポロジー表現データを更新する。そして、その内容はリンクステートルーティングプロトコル手段1500における定期処理または大きな変化があったときに、全

ての論理ノードに送出される。その後も性能評価手段1800はネットワーク内の性能評価を行っており、上述した最適化処理が反復的に行われる。

【0112】以上本発明を幾つかの実施の形態を挙げて説明したが、本発明は上記の実施の形態にのみ限定されず、その他各種の付加変更が可能であり、例えば以下のようにすることもできる。

【0113】第6発明乃至第9発明では、下位レイヤのトポロジー情報を上位レイヤにどのように見せるかという部分だけのAggregationを扱ったが、上位レイヤ内では同じように論理リンクが複数あるため、第1乃至第5発明に示される同一レイヤ内でのリンクのAggregation技術を第6乃至第9発明に対して適用することができる。

【0114】ネットワークの性能評価項目として、呼損率、ネットワーク利用率、経路情報量を用いたが、他の項目を用いることもできる。他の項目としては、例えば再経路探索回数(Crankback頻度)がある。或るAggregateリンクあるいはネットワークを利用してコネクションを設立しようとして失敗した場合、他の経路を再び検索するが、その回数が再経路探索回数である。つまり、試みたリンクでコネクションが確率できない場合に、何度他の経路を調べたかを示す。余りにAggressiveにリソースをFloodingすると、頻繁にコネクション設定が失敗し、その結果再経路探索回数が増える傾向にあるため、呼損率と同様に扱うことが可能である。

【0115】

【発明の効果】以上説明したように、本発明のトポロジー情報交換装置は、以下の効果を有する。

【0116】隣接ノード間で要約尺度を交換し、その要約尺度に基づきリンク情報を要約するため、隣接ノードA-B間の同一リンクに対してA側の要約度合いとB側の要約度合いとを一致させることができる。これにより、同じリンクでありながら、表現精度が不統一になってしまうことによる問題を防止しつつ、ネットワークにおけるノード間でリンク情報を効果的に要約することができる。

【0117】ネットワークの性能評価結果に基づき要約尺度を求め直すことにより、要約尺度の最適化が図れる。

【0118】隣接ノード間の複数リンクを類似した品質ごとにグループ化し、各リンクグループ毎に要約リンク情報を計算することにより、要約精度を或る程度保証しつつ、経路情報量を低減することができる。

【0119】ネットワークの性能評価結果に基づきリンクグループの再構成を行うことにより、リンクグループの最適化が図れる。

【0120】階層ネットワークにおいて、上位ノードが自ノード配下の下位ドメインのトポロジー情報を代表的

に交換するトポロジー情報交換装置において、下位ドメイン内のノードのうち他のドメインと接続している境界ノードを検出し、当該検出境界ノード間にあるトポロジーテンプレートで表現し、実際のトポロジーとネットワークの品質とを当該トポロジーテンプレートにより近似表現してノード間で新規要約トポロジー情報をお互いに交換することにより、大規模な階層的なネットワークにおいて、上位階層のノード間で自ノード配下の下位ドメインのトポロジー情報の効果的な要約が可能となり、ネットワーク全体としての経路情報交換量の削減が可能になる。また、効果的に内部のトポロジー情報を要約して交換することにより、ドメイン内部のある程度の荒いトポロジー情報、あるいはQOS情報をしることができるため、ドメイン間の経路選択において、どのドメインを経由すべきという最適経路選択ができるかぎり可能となり、コネクション設定失敗(ブロック)率等を低減することができる。

【0121】下位ドメイン間の複数リンクの一つあるいは少数の上位ノード間のリンクの要約に伴って境界ノードも要約することにより、経路情報量のより一層の削減が図れる。

【0122】上位ノード間で要約尺度を交換し、この要約尺度に基づき下位ドメインのトポロジー情報を計算して上位ノード間で交換することにより、各ドメインの要約精度のバラツキを防止することができる。

【0123】ネットワークの性能評価結果に基づき要約尺度を求め直すことで、下位ドメインのトポロジー情報の要約尺度の最適化が図れる。

【0124】以上の結果、大規模な階層的なネットワークにおいて、各階層内のノード間、あるいは階層間のノード間でリンク、あるいはノードのトポロジー情報、あるいはコネクション品質(クオリティオブサービス(Quality of Service; QOS))情報を経路情報としてお互いに交換して、ネットワーク全体のトポロジー情報を認識し、経路選択を行う場合、特にあるドメイン(Peer Group)のリンク、あるいはノードのトポロジー情報を全てのドメインに提供すると、経路情報が過多となり、大規模網に提供できないが、本発明によれば、効果的に内部のトポロジー情報を要約することにより、ネットワーク全体としての経路情報交換量の削減をすることが可能である。

【図面の簡単な説明】

【図1】本発明のトポロジー情報交換装置を適用したネットワークの簡単な例を示す図である。

【図2】第1の発明にかかるトポロジー情報交換装置の実施の形態の機能構成を示す図である。

【図3】Hello Packetでやりとりされる情報の例を示す図である。

【図4】隣接ノード間リンク情報の内容例を示す図である。

【図5】Aggregationパラメータ手段が隣接ノード間で行うパラメータのNegotiationの方法の例を示す図である。

【図6】自ノードリンクステートデータベースに格納されるデータの例を示す図である。

【図7】リンク情報Aggregation手段が、複数の物理リンクのQOSをどのように1本の論理リンクのQOSとして表現するのかの例を示す図である。

【図8】リンク属性の説明図である。

【図9】複数の物理リンクを論理リンクに要約する1段階目の手順の説明図である。

【図10】複数の物理リンクを論理リンクに要約する2段階目の手順の説明図である。

【図11】第2の発明にかかるトポロジー情報交換装置の実施の形態の機能構成を示す図である。

【図12】AggregationパラメータとAggregateリンクにおける呼損率およびネットワーク利用率との関係を示す図である。

【図13】図7のAggregationパラメータ交換手段が隣接ノード間で行うパラメータのNegotiationの方法の例を示す図である。

【図14】第3の発明にかかるトポロジー情報交換装置の実施の形態の機能構成を示す図である。

【図15】Aggregationグループ計算手段によるリンクのグループ化の説明図である。

【図16】Aggregationグループ計算手段が管理するデータの説明図である。

【図17】自ノード用リンクステートデータベースに格納されるデータの例を示す図である。

【図18】第4および第5の発明にかかるトポロジー情報交換装置の実施の形態の機能構成を示す図である。

【図19】リンクグループ毎のAggregateリンクが何であるかを管理しておくデータの説明図である。

【図20】階層的なPNNIルーティングのドメイン(ピアグループ; Peer Group)構成を示す図である。

【図21】論理グループノードのハローバケットのフォーマット例を示す図である。

【図22】上位Peer Group用隣接ノード間リンク情報の一例を示す図である。

【図23】第6および第7の発明にかかるトポロジー情報交換装置の実施の形態の機能構成を示す図である。

【図24】中間形式作成手段およびTopology Aggregation計算手段の処理説明図である。

【図25】中間形式作成手段の処理説明図である。

【図26】Border Node Aggregation手段の処理説明図である。

【図27】第8の発明にかかるトポロジー情報交換装置の実施の形態の機能構成を示す図である。

【図28】Aggregationパラメータの例を示

す図である。

【図29】第9の発明にかかるトポロジー情報交換装置の実施の形態の機能構成を示す図である。

【図30】性能評価手段がAggregateトポロジーが何であるかを認識するために管理するデータの例を示す図である。

【符号の説明】

100, 101, 102, 300, 301, 302, 303, 304, 305, 306, 307, 1000, 1001…ATM Node

110, 111, 112…トポロジー情報交換装置

120, 121, 122…ATM Switch Hardware

130, 131, 132, 310, 311…Node間リンク

200, 1500…リンクステートルーティングプロトコル手段

201…他ノード用リンクステートデータベース

202…自ノード用リンクステートデータベース

210, 1510…Helloプロトコル手段

211…隣接ノード間リンク情報

220, 1700…要約尺度(Aggregationパラメータ)交換手段

221, 902…リンク情報要約(Aggregation)手段

290, 1590…他ノードとのルーティング情報交換

291, 1591…隣接ノードとのリンク情報交換

292, 910, 1701…隣接ノードとの間で交換

320, 322…ネットワーク内トポロジー情報(Topology Database)

321, 323…隣接リンク情報(Neighbor Link Database)

700, 1800…性能評価手段

701, 1801…Aggregationパラメータ計算手段

900…Aggregationグループ通知手段

901…リンクグループ作成手段

903…Aggregationグループ計算手段

1200, 1201, 1202, 1310, 1311,

1312, 1313…Aggregation Parameter Negotiation手順

1400, 1401, 1402…最下層Peer Group

1404…上位階層Peer Group

1410, 1411, 1412…上位階層論理ノード

1420, 1421, 1422, 1423…Border Node

1440, 1441, 1442, 1443, 1444…最下層ノード間リンク

1450, 1451, 1452, 1453…上位階層ノ

ード間論理リンク

1470, 1471, 1472...Virtual Circuit

1480...最下層代表ノード (Peer Group Leader)

1490...Topology Aggregation (Border Node Aggregationなし)

1491...Topology Aggregation (Border Node Aggregationあり) 10

1501...下位Peer Group用リンクステートデータベース

1502...上位Peer Group用リンクステートデータベース

1511...下位Peer Group用隣接ノード間リンク情報

* 1512...上位Peer Group用隣接ノード間リンク情報

1520...下位Peer Group Border Node検出手段

1521...中間形式作成手段

1522...Topology Aggregation 計算手段

1523...Border Node Aggregation手段

1601...Peer Group内物理トポロジー

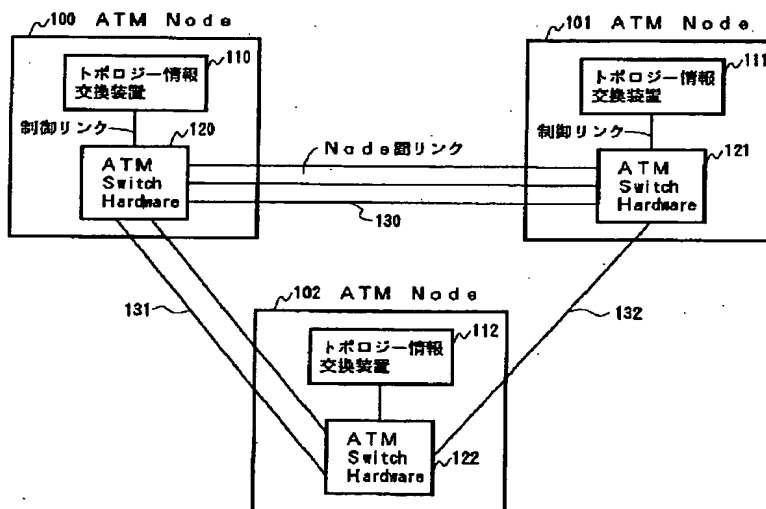
1602...中間形式

1603...Symmetric StarでのAggregation

1604...Assymmetric Star (Exceptionなし)でのAggregation

* 1605...Assymmetric Star (Exceptionつき)でのAggregation

【図1】



【図3】

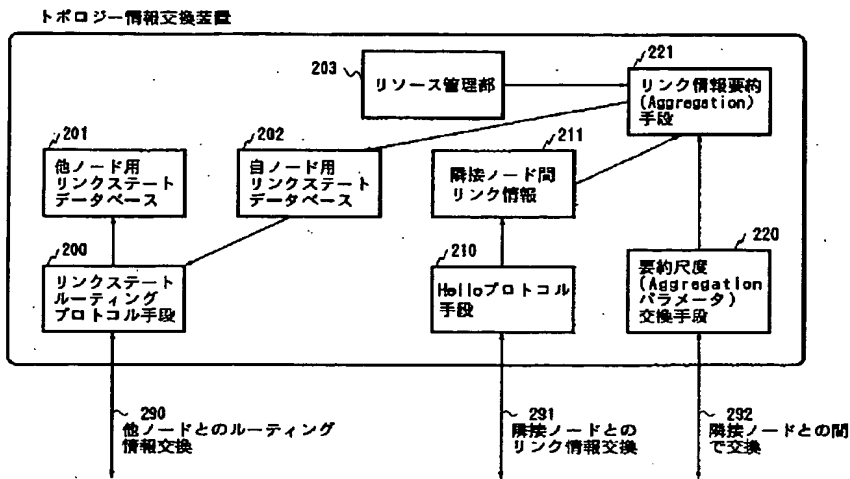
My Node ID
My Port ID
Remote Node ID
Remote Port ID
Peer Group ID (Domain ID)

【図8】

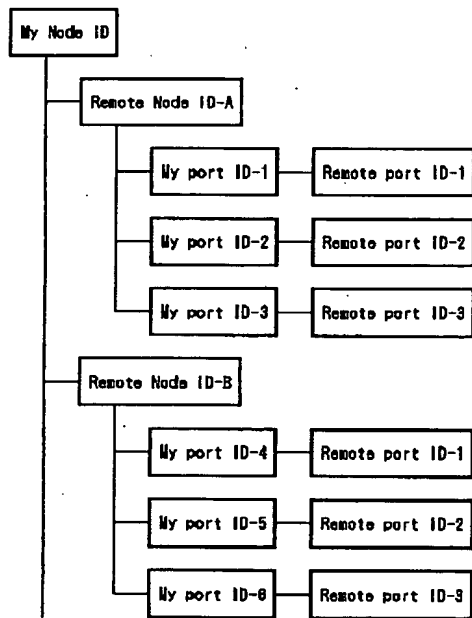
Link属性

トラフィッククラス	CBR	rt-VBR	nrt-VBR	ABR	UBR
QOS metrics	Maximum Cell Rate (MaxCR)	Available Cell Rate (AvCR)	Cell Delay Variation (CDV)	Cell Transfer Delay (CTD)	Cell Loss Ratio (CLR)
他の metrics	Administrative Weight (AW)				

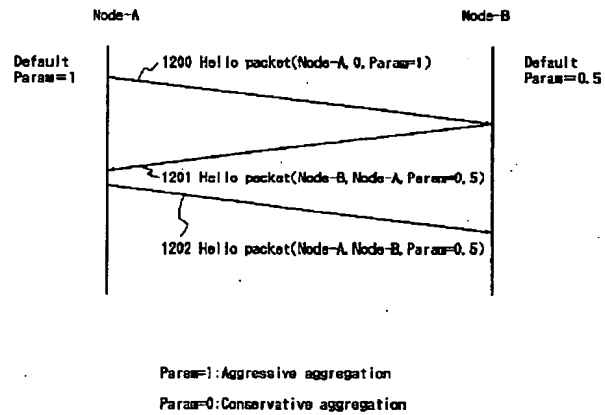
【図2】



【図4】



【図5】

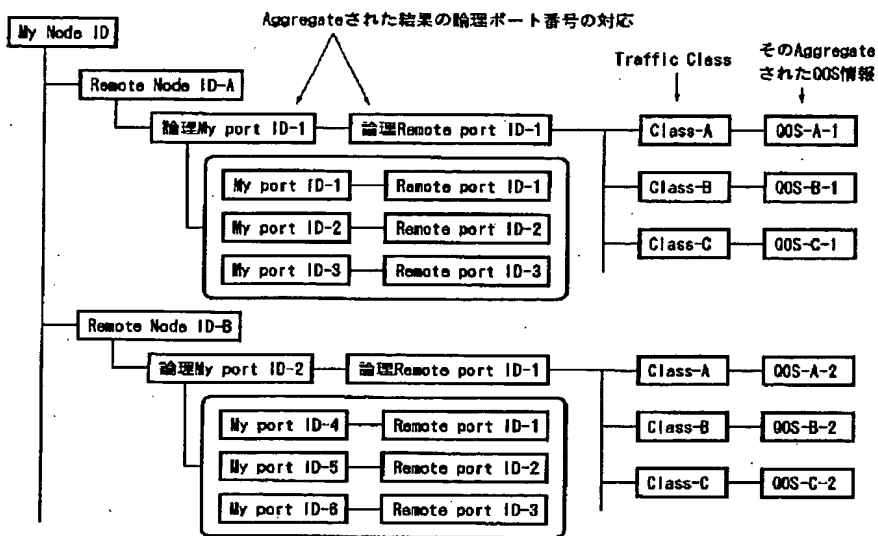


【図9】

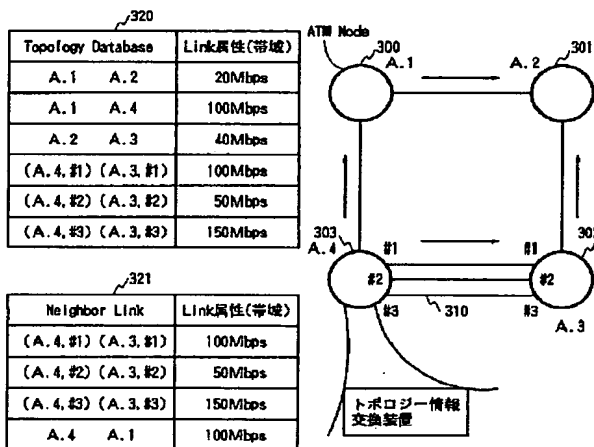
Link-1	Class-A QOS-1	Class-B QOS-2		
Link-2	Class-A QOS-3		Class-C QOS-4	Class-D QOS-5
Link-3	Class-A QOS-6	Class-B QOS-7	Class-C QOS-8	
Aggregated Link	Class-A QOS-1 QOS-3 QOS-6	Class-B QOS-2 QOS-7	Class-C QOS-4 QOS-8	Class-D QOS-5

QOS-N(N=1,...8)はQOS metrics vectorから構成される

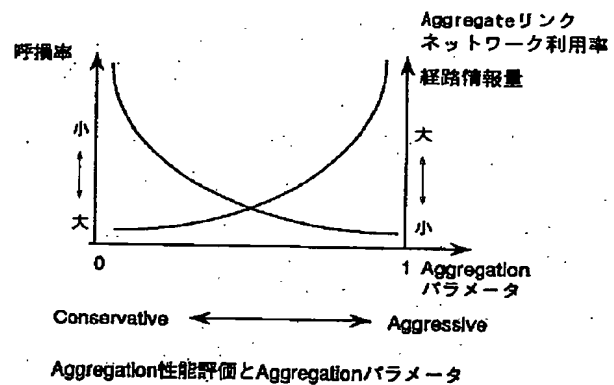
【図6】



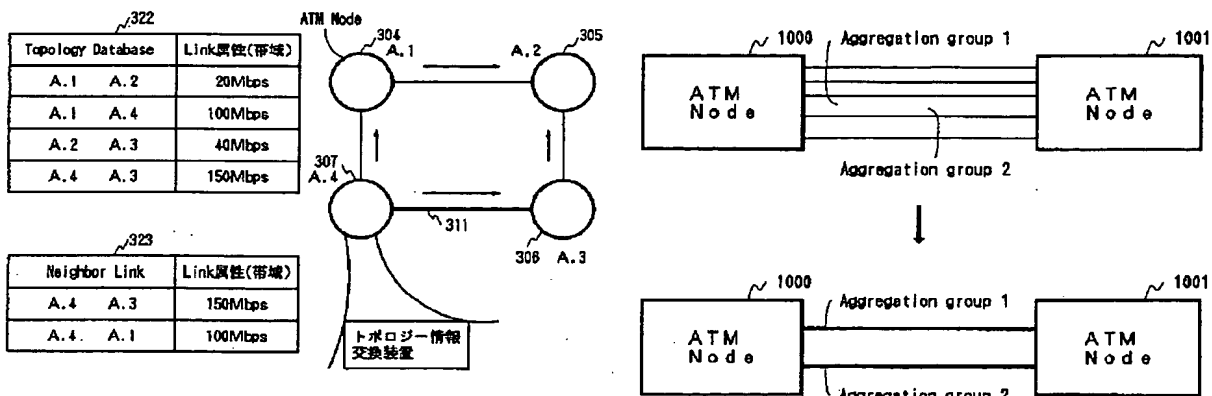
【図7】



【図12】



【図15】



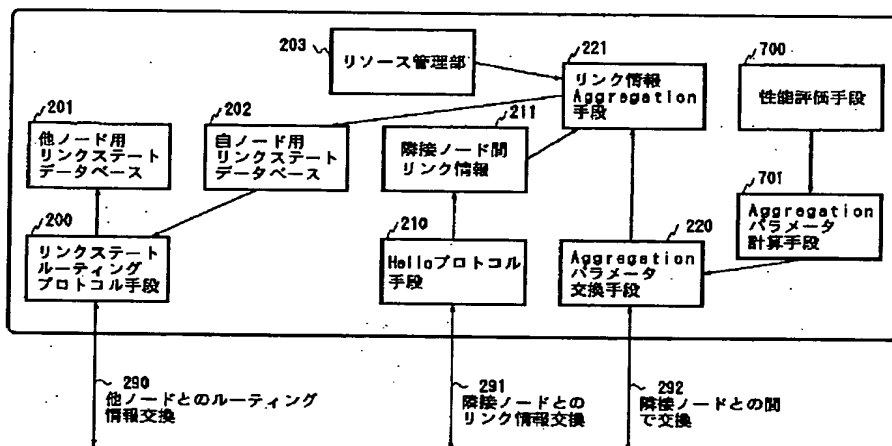
【図10】

	MaxCR (Mbps)	AvCR (Mbps)	CDV (usec)	CTD (usec)	CLR	AW
QOS1	155	100	10	100	10^{-6}	10
QOS2	155	10	20	200	10^{-5}	100
QOS3	100	80	30	250	10^{-6}	20
Aggressive Aggregation ($r=1$)	Max 155	Max 100	Min 10	Min 100	Min 10^{-6}	Min 10
Conservative Aggregation ($r=0$)	Min 100	Min 10	Max 30	Max 250	Max 10^{-5}	Max 100
中間的 Aggregation ($r=r$)	$r \cdot 155 + (1-r) \cdot 100$	$r \cdot 100 + (1-r) \cdot 10$	$r \cdot 10 + (1-r) \cdot 30$	$r \cdot 100 + (1-r) \cdot 250$	$r \cdot 10^{-6} + (1-r) \cdot 10^{-5}$	$r \cdot 10 + (1-r) \cdot 100$

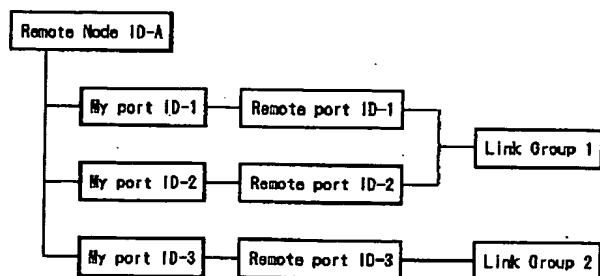
【図28】

Aggregation method (Aggressive, or Conservative)	
収束条件	1. バイパスリンク最大数
	2. 近似精度閾値

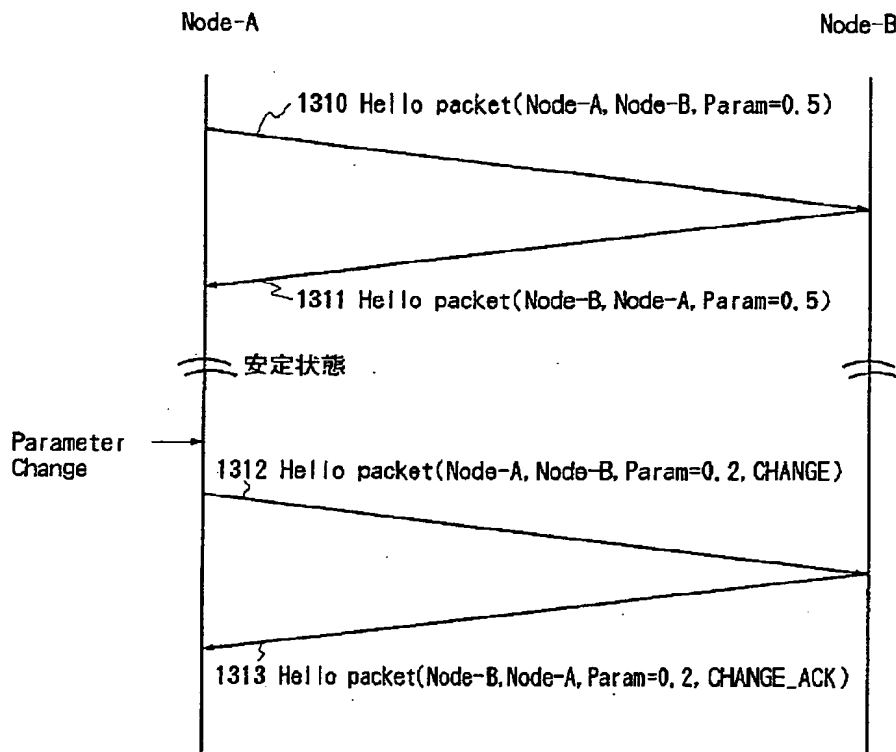
【図11】



【図16】



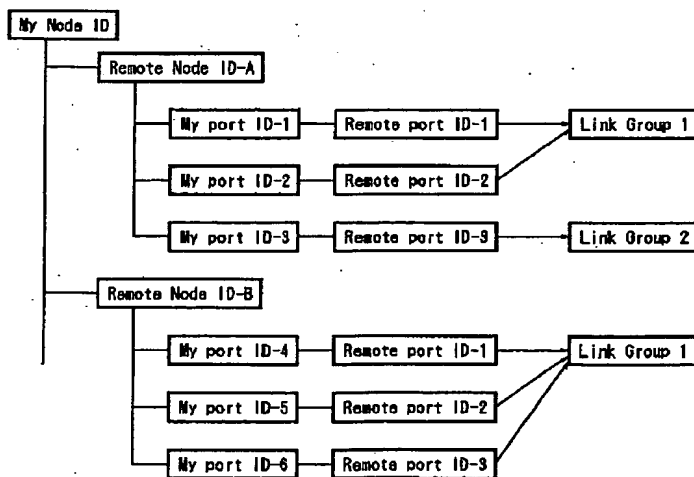
【図 13】



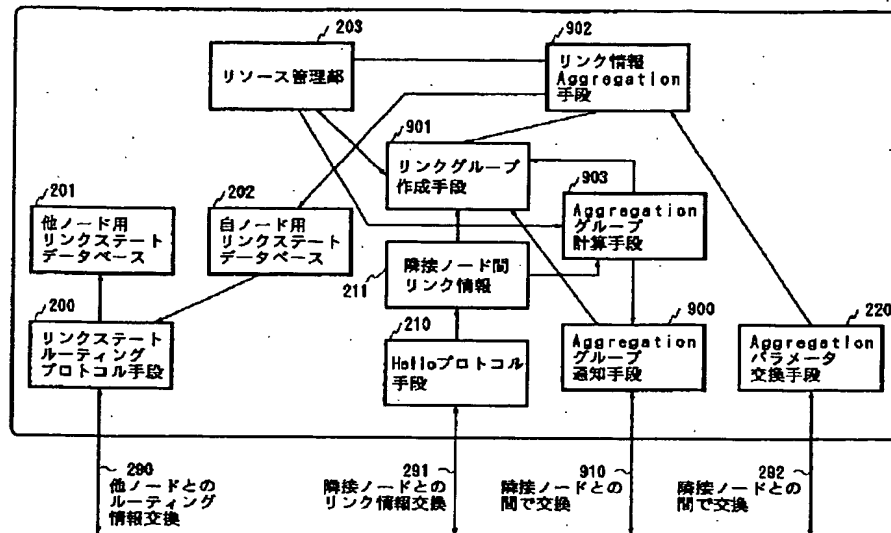
Param=1:Aggressive aggregation

Param=0:Conservative aggregation

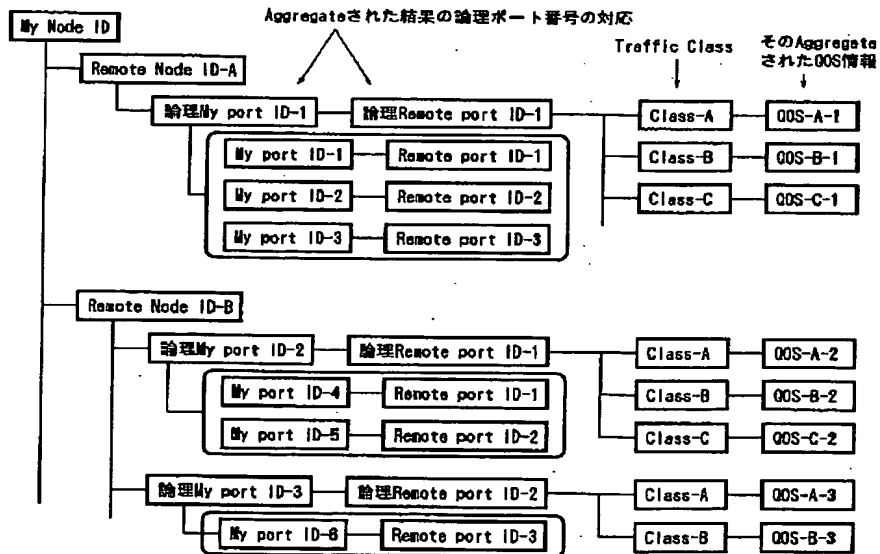
【図 19】



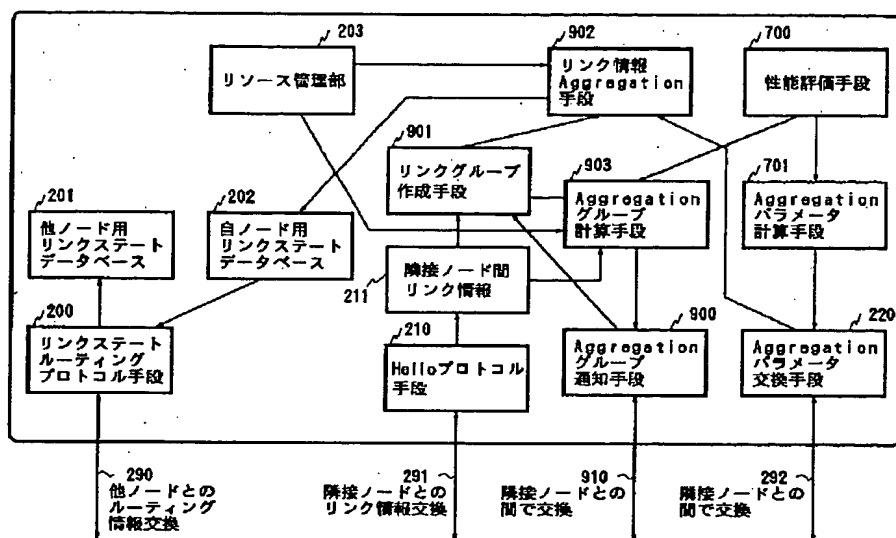
【図14】



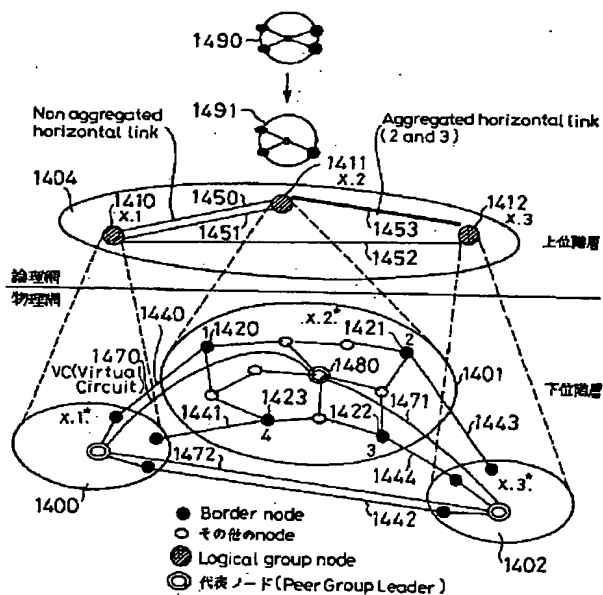
【図17】



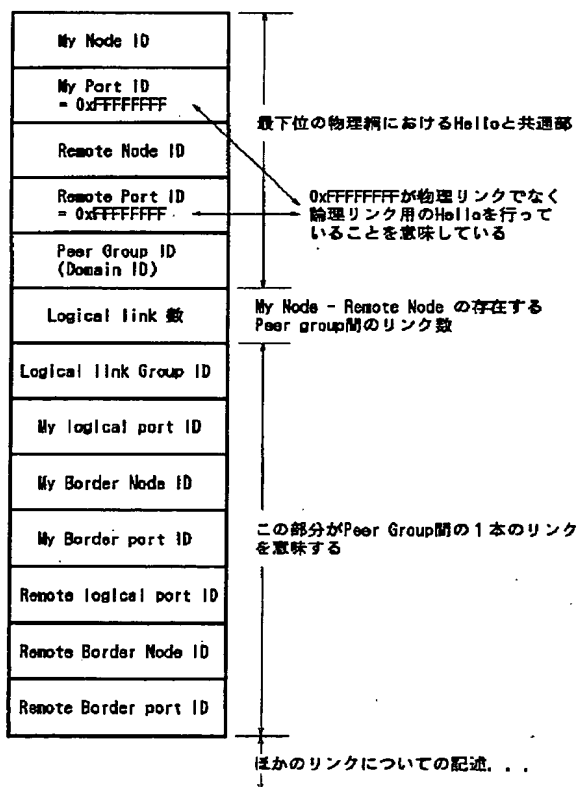
【図18】



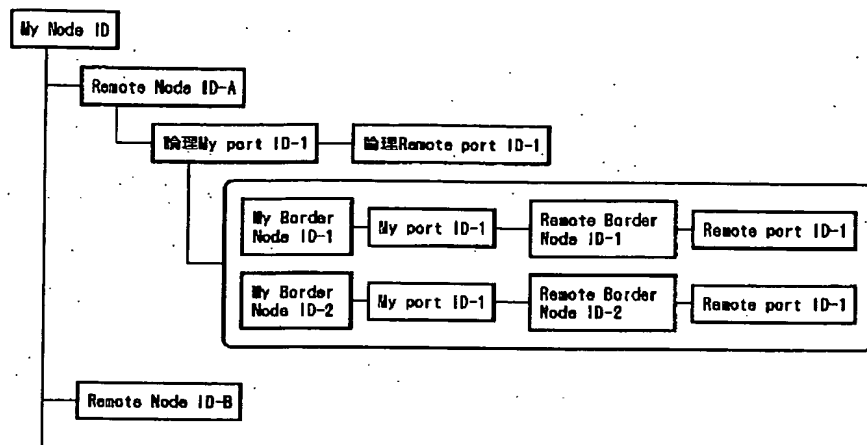
【図20】



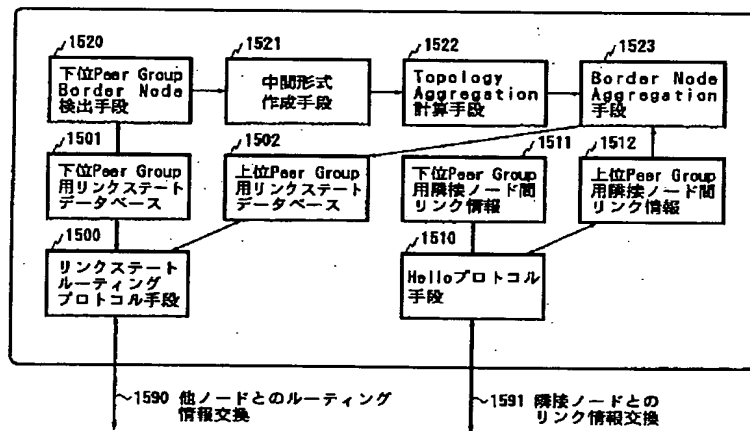
【図21】



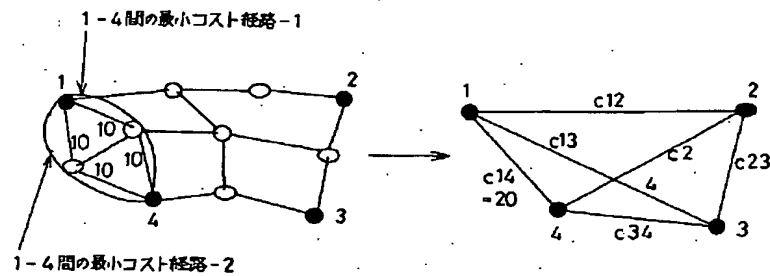
【図22】



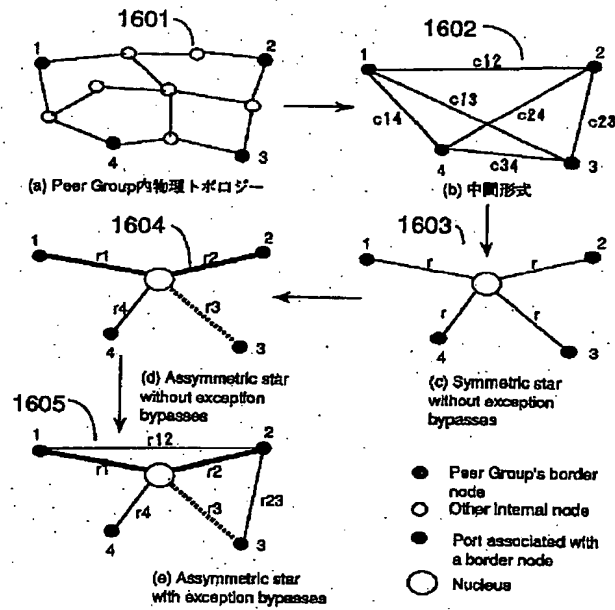
【図23】



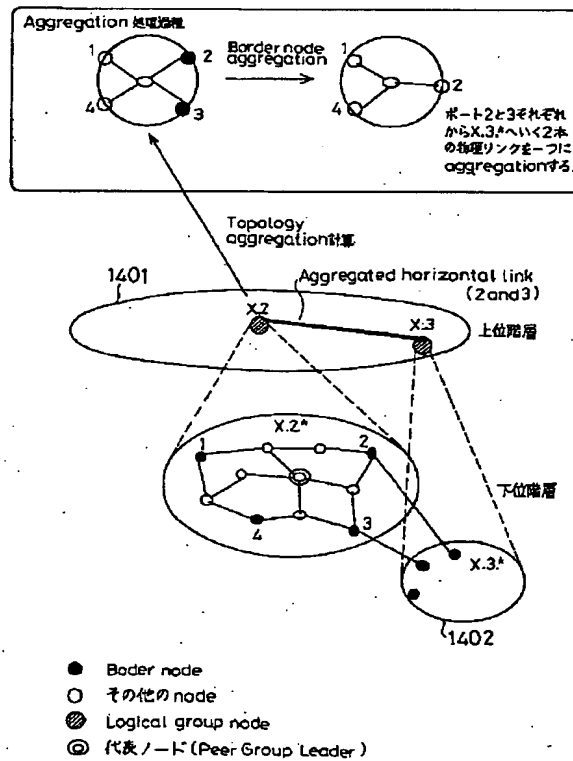
【図25】



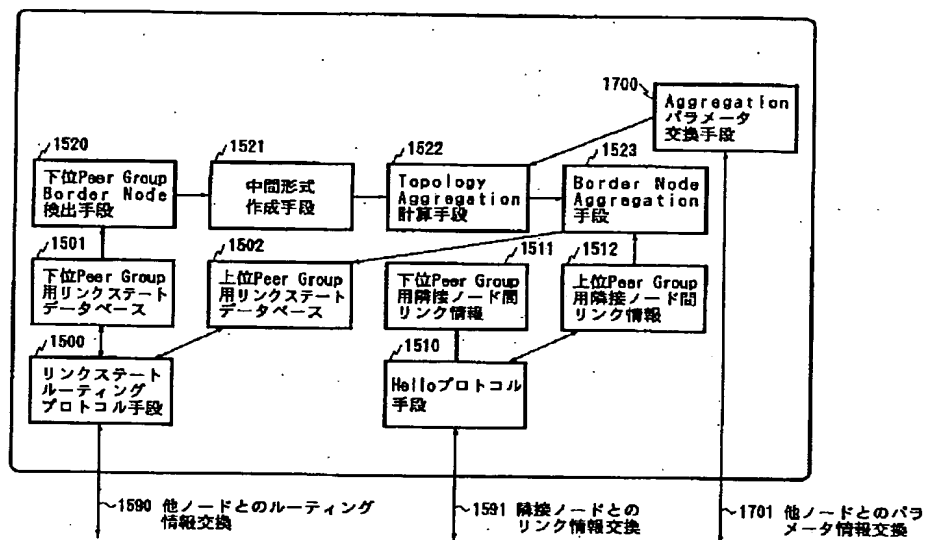
【図24】



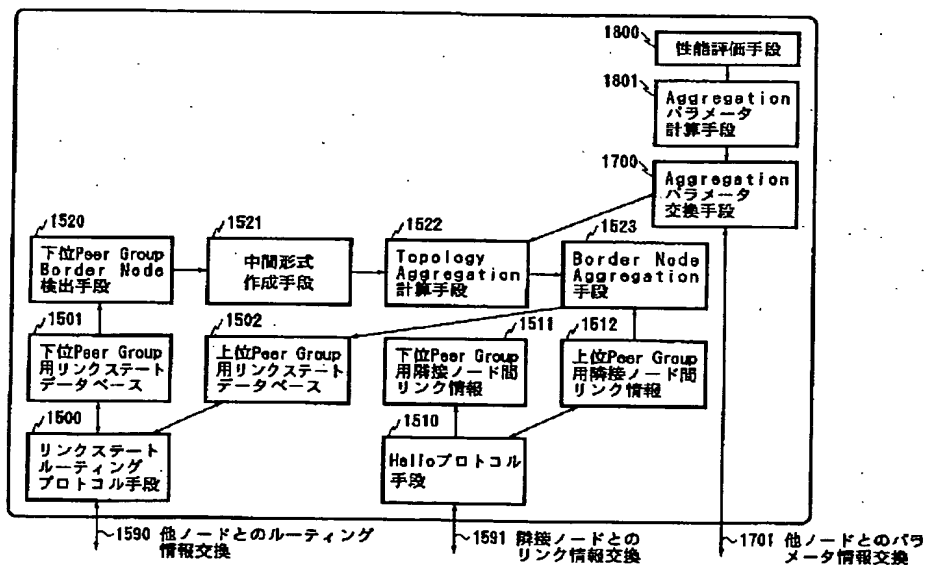
【図26】



【図27】



【図29】



【図 30】

